

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2005, 20:03:24 ; Search time 166 Seconds
(without alignments)
766.531 Million cell updates/sec

Title: us-10-687-268-35

Perfect score: 1756

Sequence: 1 MVTKAFVLLAIPAEASAKSC.....YFSDVLHLLIMLIGICAVL 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1756	100.0	329	5	Aae17314 Human tec
2	1756	100.0	329	5	Adh48858 NOV60C pr
3	1756	100.0	329	8	Adh72158 Human pro
4	1751	99.7	329	5	Adh48856 NOV60B pr
5	1751	99.7	329	8	Adh72162 Human pro
6	1539.5	87.7	300	5	Adh48854 NOV60A pr
7	1539.5	87.7	300	8	Adh72160 Human pro
8	1505	85.7	276	8	Adh72164 Human pro
9	1478	84.2	273	8	Adh72166 Human pro
10	248	14.1	415	5	Abg72418 Cytochrom
11	228.5	13.0	644	8	Ado59937 Rat UMOD
12	223.5	12.7	642	8	Ado59936 Murine UM
13	222	12.6	621	8	Abm82705 Human dia
14	221	12.6	640	8	Abm82704 Human dia
15	220	12.5	643	8	Ado59935 Bovine UM
16	219	12.5	507	8	Adr09416 Human pro
17	219	12.5	577	7	Adm04800 Human pro
18	219	12.5	673	7	Adb64620 Human pro
19	213	12.1	406	3	Aab54274 Human pan
20	210	12.0	527	8	Adi02916 Human pan
21	210	12.0	530	5	Aae17559 Human pan
22	210	12.0	534	6	AbR39945 Human pro
23	210	12.0	640	8	Ado59934 Human UMO
24	210	12.0	640	8	Adt49902 Human UMO
25	209	11.9	384	2	Aar98963 Human GP2

ALIGNMENTS

RESULT 1

AAE17314
ID AAE17314 standard; protein; 329 AA.

XX AC AAE17314;

XX AC AAE17314;

DT 18-APR-2002 (first entry)

DE Human tectorin beta protein, sbg453915TECTORINA.

XX Human; therapy; wound healing disorder; vaccine; cancer; infection;
KW autoimmunity disorder; haematopoietic disorder; inflammation; arthritis;
KW Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic;
KW multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;
KW ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;
KW depression; cardiovascular disease; myocardial infarction; renal failure;
KW respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;
KW type II diabetes mellitus; skeletal muscle disease; immunosuppressive;
KW hyperplasia; renal disease; hypoglycaemia; gastrointestinal disease;
KW neurotic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;
KW haemostatic; vulnery; anticonvulsant; antirheumatic; neuroprotective;
KW nephrotropic; hypotensive; vasotropic; cytostatic; cerebroprotective;
KW allergy; tectorin beta.

XX Homo sapiens.

XX WO200198342-A1.

XX 27-DEC-2001.

XX 22-JUN-2001; 2001WO-US019929.

XX 22-JUN-2000; 2000US-0213156P.

XX 22-JUN-2000; 2000US-0213161P.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX (GLAX) GLAXO GROUP LTD.

XX Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA;

XX Murdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;

XX WPI; 2002-139783/18.

XX N-FSDB; AAD27809.

XX Novel secreted and membrane-associated polypeptides and polynucleotides
XX useful for preventing, ameliorating or correcting dysfunction or disease
XX including diabetes, cancer, hypertension and growth abnormalities.

Brannock, M,
101687268
Seq. ID 35

XX	PS	Claim 1; Page 127-128; 138pp; English.	DE	XX	NOV60C protein sequence, SEQ ID 142.
XX	CC	The invention relates to secreted and membrane-associated polypeptides and polynucleotides. The sequences of the invention are useful in diagnostic assays for detecting diseases associated with inappropriate activity or levels of these polynucleotides, and in identifying their agonists and antagonists that are potentially useful in therapy. The sequences of the invention are useful as vaccines for inducing an immunological response. The sequences of the invention are useful for treating cancers, infections, autoimmune disorders, haematopoietic disorders, wound healing disorders, cholesterol ester storage disease, inflammation, congenital muscular dystrophy, junctional epidermolysis bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis, viral and bacterial infections, Alzheimer's disease, asthma, arthritis, allergies, schizophrenia, sbg4445PROA-associated disorders, septicemia, psoriasis, inflammatory bowel disease, transplant rejection, graft versus host disease, ischaemia, stroke, acute respiratory disease, syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis, brain disorders including parasupranuclear palsy, myotonic dystrophy, depression, anxiety disorders and sleep disorders, cardiovascular diseases including congestive heart failure and myocardial infarction, respiratory diseases including chronic obstructive pulmonary disease, acute bronchitis and adult respiratory distress syndrome, liver disorders including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral and non-viral hepatitis, type II diabetes mellitus, renal disease including acute and chronic renal failure, glomerulonephritis, Fanconi's syndrome, cystinuria, skeletal muscle disorders including hypoglycaemia and tendinitis, gastrointestinal diseases including intestinal obstruction and tropical sprue, spleen disorders including hypersplenism, Hodgkin's disease and malignant lymphoma, testicular cancer, male reproductive diseases including low testosterone and male infertility. The present sequence is human tectorin beta protein	XX	Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic; hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV60C; beta tectorin-like protein; chromosome 10.	
XX	CC	Sequence 329 AA;	XX	OS	Homo sapiens.
		Query Match 100.0%; Score 1756; DB 5; Length 329;	XX	XX	WO200268652-A2.
		Best Local Similarity 100.0%; Pred. No. 1.8e-180;	XX	XX	06-SEP-2002.
		Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	XX	26-FEB-2002; 2002WO-US005910.
QY	1	MTYKAFVLLAIFAEASAKSCAPNKADVLVFCYPTKIITKIPECYGVHQLALGGLCY 60	XX	XX	26-FEB-2001; 2001US-0271646P.
Db	1	MTYKAFVLLAIFAEASAKSCAPNKADVLVFCYPTKIITKIPECYGVHQLALGGLCY 60	XX	XX	27-FEB-2001; 2001US-0271840P.
QY	61	NGVHEGGYQFVLPDLPKNSYCGTQSEYKPPYHVFYSHIVSNDTIVIKNQPNYSFS 120	XX	XX	28-FEB-2001; 2001US-0272404P.
Db	61	NGVHEGGYQFVLPDLPKNSYCGTQSEYKPPYHVFYSHIVSNDTIVIKNQPNYSFS 120	XX	XX	28-FEB-2001; 2001US-0272405P.
QY	121	CTYHSTYLVNQAFDQVAVTHVKNMGCTPESQSLNFTYTNKFSIKKEAPFVLEAEI 180	XX	XX	28-FEB-2001; 2001US-0272410P.
Db	121	CTYHSTYLVNQAFDQVAVTHVKNMGCTPESQSLNFTYTNKFSIKKEAPFVLEAEI 180	XX	XX	28-FEB-2001; 2001US-0272414P.
QY	181	GSDLPAGVEAKGLSIRFKVVLNSCWATPSADPMYPLQWLINKGCTPTDTVLVHNGRDH 240	XX	XX	02-MAR-2001; 2001US-0272787P.
Db	181	GSDLPAGVEAKGLSIRFKVVLNSCWATPSADPMYPLQWLINKGCTPTDTVLVHNGRDH 240	XX	XX	02-MAR-2001; 2001US-0272922P.
QY	241	RATPFQNAFPNIPKLSKVWLHCETFCIDSEKLSCPVTCDDRKLLRDQTGGVLVWELS 300	XX	XX	02-MAR-2001; 2001US-0273048P.
Db	241	RATPFQNAFPNIPKLSKVWLHCETFCIDSEKLSCPVTCDDRKLLRDQTGGVLVWELS 300	XX	XX	02-MAR-2001; 2001US-0273300P.
QY	301	LRSRGFSSLYSFSFVLLHLMGLICAVL 329	XX	XX	16-MAR-2001; 2001US-0273401P.
Db	301	LRSRGFSSLYSFSFVLLHLMGLICAVL 329	XX	XX	20-MAR-2001; 2001US-0277324P.
		RESULT 2	XX	XX	30-MAR-2001; 2001US-0278660P.
		ADH48858	XX	XX	30-MAR-2001; 2001US-0280039P.
XX	ID	ADH48858 standard; protein; 329 AA.	XX	XX	30-MAR-2001; 2001US-0280234P.
XX	AC	ADH48858;	XX	XX	12-APR-2001; 2001US-0280818P.
XX	DT	25-MAR-2004 (first entry)	XX	XX	12-APR-2001; 2001US-0283443P.
XX			XX	XX	23-APR-2001; 2001US-0285754P.
			XX	XX	24-APR-2001; 2001US-0286096P.
			XX	XX	03-MAY-2001; 2001US-0288353P.
			XX	XX	17-MAY-2001; 2001US-0291703P.
			XX	XX	31-MAY-2001; 2001US-0294834P.
			XX	XX	20-JUN-2001; 2001US-0296959P.
			XX	XX	21-JUN-2001; 2001US-0299845P.
			XX	XX	05-JUL-2001; 2001US-0303242P.
			XX	XX	13-AUG-2001; 2001US-0311981P.
			XX	XX	16-AUG-2001; 2001US-0312858P.
			XX	XX	17-AUG-2001; 2001US-0313280P.
			XX	XX	29-AUG-2001; 2001US-0315614P.
			XX	XX	17-SEP-2001; 2001US-0322818P.
			XX	XX	25-FEB-2002; 2002US-00322818.
			XX	XX	(CURA-) CURAGEN CORP.
			XX	XX	Alsobrook JP, Anderson DW, Ballinger RA, Boldog FL, Burgess CE; Casman SJ, Ellerman KE, Gangolli EA, Gerlach VL, Gilbert JA; Gorman L, Guo X, Gusev VY, Kikuda R, Li L, Liu X, Malyankar UM; Miller CE, Millet I, Padigaru M, Patturajan M, Pena CEA, Feyman JA; Rastelli L, Shenoy SG, Shimkets RA, Smithson G, Spytek KA, Stone DJ; Taupier RJ, Tchernev VT, Vernet CAM, Zerhusen BD;
			XX	XX	WPI; 2002-698672/75.
			XX	XX	N-FSDB; ADH48857.
			XX	XX	New NOVX polypeptides or polynucleotides, useful for preventing or treating disorders or syndromes e.g., atherosclerosis, hypertension, obesity or cancer.
			XX	XX	Claim 1; Page 338; 923pp; English.
			XX	XX	The present invention relates to novel human NOVX proteins, where X is any number from 1 to 91 and their coding sequences. The proteins and coding sequences are useful for preventing or treating disorders or syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOV60 is a beta tectorin-like protein and its coding sequence maps to chromosome 10.
			XX	XX	Sequence 329 AA;

Query Match		100.0%;	Score 1756;	DB 5;	Length 329;
-Best Local Similarity		100.0%;	Pred. No. 1.8e-180;		
Matches 329;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTKAPVLLAI	FAEASAKSCAPNKADVILVFCYPKTII	KIPECYPGWVHQLALGGLCY	60
DB	1	MTKAPVLLAI	FAEASAKSCAPNKADVILVFCYPKTII	KIPECYPGWVHQLALGGLCY	60
QY	61	NGVHEGGYQFV	IPDLSPNKKSYCGTQSYKPKPIVHYFYS	HVSNDDTVIVKQNPVNSFS	120
DB	61	NGVHEGGYQFV	IPDLSPNKKSYCGTQSYKPKPIVHYFYS	HVSNDDTVIVKQNPVNSFS	120
QY	121	CTYHSTYLVNQ	AADQORVATVHVKGSMGTFFESQLS	LNFTYTNKPSIKKEAPFVLEASI	180
DB	121	CTYHSTYLVNQ	AADQORVATVHVKGSMGTFFESQLS	LNFTYTNKPSIKKEAPFVLEASI	180
QY	181	GSDLFAGV	EAKGLSTRFKVILNSCWATPSADFMYP	LOWQLINNGCPTDTEVLVHNGRDH	240
DB	181	GSDLFAGV	EAKGLSTRFKVILNSCWATPSADFMYP	LOWQLINNGCPTDTEVLVHNGRDH	240
QY	241	RATFQNAFR	FNQIPKLSKVWLHCETFCIDSEKLS	CPVTCDKRKRLRRDQGGVLVVELS	300
DB	241	RATFQNAFR	FNQIPKLSKVWLHCETFCIDSEKLS	CPVTCDKRKRLRRDQGGVLVVELS	300
QY	301	LRSGFSSLYS	FDVHLHLMGICAVL	329	
DB	301	LRSGFSSLYS	FDVHLHLMGICAVL	329	
RESULT 3					
ID	ADH72158	standard; protein; 329 AA.			
XX	AC	ADH72158;			
XX	DT	25-MAR-2004 (first entry)			
XX	DE	Human protein of the invention NOV48a SEQ ID NO:1054.			
XX	KW	human; cytostatic; immunomodulator; neuroprotective; nootropic;			
XX	KW	anorectic; antidabetic; antimicrobial; antilipemic; gene therapy;			
XX	KW	vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;			
XX	KW	obesity; diabetes; infectious disease; metabolic syndrome X;			
XX	OS	dyslipidaemia.			
XX	PN	Homo sapiens.			
XX	PD	WO2003102155-A2.			
XX	PD	11-DEC-2003.			
XX	PF	03-JUN-2003; 2003WO-US017430.			
XX	PR	03-JUN-2002; 2002US-0385120P.			
XX	PR	04-JUN-2002; 2002US-0385784P.			
XX	PR	05-JUN-2002; 2002US-0385041P.			
XX	PR	05-JUN-2002; 2002US-0385047P.			
XX	PR	06-JUN-2002; 2002US-0385376P.			
XX	PR	06-JUN-2002; 2002US-0385453P.			
XX	PR	06-JUN-2002; 2002US-0386864P.			
XX	PR	06-JUN-2002; 2002US-0387016P.			
XX	PR	07-JUN-2002; 2002US-0386796P.			
XX	PR	07-JUN-2002; 2002US-0386816P.			
XX	PR	07-JUN-2002; 2002US-0386931P.			
XX	PR	07-JUN-2002; 2002US-0386942P.			
XX	PR	07-JUN-2002; 2002US-0386971P.			
XX	PR	07-JUN-2002; 2002US-0387262P.			
XX	PR	08-JUN-2002; 2002US-0296960P.			
XX	PR	10-JUN-2002; 2002US-0387400P.			
XX	PR	10-JUN-2002; 2002US-0387535P.			
XX	PR	11-JUN-2002; 2002US-0387610P.			
XX	PR	11-JUN-2002; 2002US-0387625P.			
XX	PR	11-JUN-2002; 2002US-0387634P.			

PR	11-JUN-2002;	2002US-0387668P.
PR	11-JUN-2002;	2002US-0387696P.
PR	11-JUN-2002;	2002US-0387702P.
PR	11-JUN-2002;	2002US-0387836P.
PR	11-JUN-2002;	2002US-0387859P.
PR	12-JUN-2002;	2002US-0387933P.
PR	12-JUN-2002;	2002US-0387934P.
PR	12-JUN-2002;	2002US-0387960P.
PR	12-JUN-2002;	2002US-0388022P.
PR	12-JUN-2002;	2002US-0388096P.
PR	13-JUN-2002;	2002US-0389123P.
PR	14-JUN-2002;	2002US-0389118P.
PR	14-JUN-2002;	2002US-0389120P.
PR	14-JUN-2002;	2002US-0389144P.
PR	14-JUN-2002;	2002US-0389146P.
PR	17-JUN-2002;	2002US-0389729P.
PR	17-JUN-2002;	2002US-0389742P.
PR	18-JUN-2002;	2002US-0389884P.
PR	19-JUN-2002;	2002US-0390066P.
PR	19-JUN-2002;	2002US-0390209P.
PR	21-JUN-2002;	2002US-0390763P.
PR	17-JUL-2002;	2002US-0396706P.
PR	06-AUG-2002;	2002US-0401628P.
PR	09-AUG-2002;	2002US-0402156P.
PR	09-AUG-2002;	2002US-0402256P.
PR	09-AUG-2002;	2002US-0402389P.
PR	12-AUG-2002;	2002US-0402786P.
PR	12-AUG-2002;	2002US-0402816P.
PR	12-AUG-2002;	2002US-0402821P.
PR	12-AUG-2002;	2002US-0402832P.
PR	13-AUG-2002;	2002US-0403448P.
PR	13-AUG-2002;	2002US-0403459P.
PR	13-AUG-2002;	2002US-0403531P.
PR	13-AUG-2002;	2002US-0403532P.
PR	13-AUG-2002;	2002US-0403563P.
PR	15-AUG-2002;	2002US-0403617P.
PR	26-AUG-2002;	2002US-0406182P.
PR	26-AUG-2002;	2002US-0406355P.
PR	27-AUG-2002;	2002US-0406240P.
PR	12-SEP-2002;	2002US-0410084P.
PR	20-SEP-2002;	2002US-0412528P.
PR	23-SEP-2002;	2002US-0412731P.
PR	30-SEP-2002;	2002US-0414801P.
PR	30-SEP-2002;	2002US-0414839P.
PR	30-SEP-2002;	2002US-0414840P.
PR	30-SEP-2002;	2002US-0414954P.
PR	09-OCT-2002;	2002US-0417186P.
PR	09-OCT-2002;	2002US-0417406P.
PR	23-OCT-2002;	2002US-0420639P.
PR	28-OCT-2002;	2002US-0421156P.
PR	31-OCT-2002;	2002US-0422690P.
PR	01-NOV-2002;	2002US-0423130P.
PR	05-NOV-2002;	2002US-00423798.
PR	05-NOV-2002;	2002US-0423798P.
XX	12-NOV-2002;	2002US-0425453P.
PA	(CURA-) CURAGEN CORP.	
XX	Alsobrook JP,	Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI	Catterton E,	Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI	Ettenberg S,	Gerlach VL, Gorman L, Gunther E, Guo X;
PI	Gusev VY,	Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI	MacLachlan T,	Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI	Padigaru M,	Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI	Riesger DK,	Rothenberg ME, Sciore P, Shenoy SG, Shinkets RA;
PI	Smithson G,	Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
XX	Zhong H;	
DR	WPI; 2004-081935/08.	
DR	N-PSDB; ADH72157.	
XX		
PT	New NOVX polypeptides and nucleic acid molecules useful for preventing or	

Db 121 CTHSTYLVNQAFDORVATVHVKNKSGMTFFESQLSNFYTNKAFSIKKEAPFVLEASEI 180
QY 181 GSDLPAGVAKGISIRFKVVLNSCWATPSADFMYPLOWOLINKGCPDTETVLVHENGDRH 240
Db 181 GSDLPAGVAKGISIRFKVVLNSCWATPSADFMYPLOWOLINKGCPDTETVLVHENGDRH 240
QY 241 RATFQFNAPFRFQNIPLKSKVWLHCETFFICDSEKLSCPVTCDKRKLRLRDQTGGVLVVELS 300
Db 241 RATFQFNAPFRFQNIPLKSKVWLHCETFFICDSEKLSCPVTCDKRKLRLRDQTGGVLVVELS 300
QY 301 LRSRGFSSLYSFDVLHHLIMLIGICAVL 329
Db 301 LRSRGFSSLYSFDVLHHLIMLIGICAVL 329

RESULT 5

ADH72162
ID ADH72162 standard; protein; 329 AA.
XX
AC ADH72162;
XX
DT 25-MAR-2004 (first entry)
XX
DE Human protein of the invention NOV48c SEQ ID NO:1058.
XX
KW human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.

PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402158P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 12-SEP-2002; 2002US-0406240P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VV, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI MacLachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CE, Feyman AG, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
WPI: 2004-081935/08.
DR N-PSDB; ADH72161.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 48; SEQ ID NO 1058; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The

CC polypeptide, polynucleotide and antibody are useful in diagnosing.
 CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
 CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
 CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
 CC further used as hybridisation probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. The present sequence
 CC represents a NOVX polypeptide of the invention.
 XX
 - SQ Sequence 329 AA;
 Query Match 99.7%; Score 1751; DB 8; Length 329;
 Best Local Similarity 99.7%; Pred. No. 6.4e-180;
 Matches 328; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MVTAFVLLAIFAEASAKSCAPNKADVLVFCYPTKIITIKIPECPYGEVHQLALGGLCY 60
 Db 1 MVTAFVLLAIFAEASAKSCAPNKADVLVFCYPTKIITIKIPECPYGEVHQLALGGLCY 60
 QY 61 NGVHEGGYQFVPIPLSPKNKSYCGTQSEYKPPYHFSYSHIVSNDTIVVKNQPNYSFS 120
 Db 61 NGVHEGGYQFVPIPLSPKNKSYCGTQSEYKPPYHFSYSHIVSNDTIVVKNQPNYSFS 120
 QY 121 CTYHSTYLVNQAAFDQVRVATVHVKNKSGMTFESQLSLNFYTNAKFSIKKEAPFVLEA 180
 Db 121 CTYHSTYLVNQAAFDQVRVATVHVKNKSGMTFESQLSLNFYTNAKFSIKKEAPFVLEA 180
 QY 181 GSDLFAGVEAKGLSIRFKVVLNSCWATPSADPMYPLQWLKNGCPTDETIVLVHNGRDH 240
 Db 181 GSDLFAGVEAKGLSIRFKVVLNSCWATPSADPMYPLQWLKNGCPTDETIVLVHNGRDH 240
 QY 241 RATFQFNAPRFONIPKLSKVLHCHETFCIDSEKLSKCPVTCDKRLRLDQTGGVLVVELS 300
 Db 241 RATFQFNAPRFONIPKLSKVLHCHETFCIDSEKLSKCPVTCDKRLRLDQTGGVLVVELS 300
 QY 301 LRSRGFSSLYSFDVLHLLIMMLGICAVL 329
 Db 301 LRSRGFSSLYSFDVLHLLIMMLGICAVL 329
 RESULT 6
 ADH48854
 ID ADH48854 standard; protein; 300 AA.
 XX
 AC ADH48854;
 XX
 DT 25-MAR-2004 (first entry)
 XX
 DE NOV60A protein sequence, SEQ ID 138.
 XX
 KW Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;
 KW hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV60A;
 KW beta tectorin-like protein; chromosome 10.
 XX
 OS Homo sapiens.
 XX
 PN WO200268652-A2.
 XX
 PD 06-SEP-2002.
 XX
 PF 26-FEB-2002; 2002WO-005910.
 XX
 PR 26-FEB-2001; 2001US-0271646P.
 PR 27-FEB-2001; 2001US-0271840P.
 PR 28-FEB-2001; 2001US-0272404P.
 PR 28-FEB-2001; 2001US-0272405P.
 PR 28-FEB-2001; 2001US-0272410P.
 PR 28-FEB-2001; 2001US-0272414P.
 PR 02-MAR-2001; 2001US-0272787P.
 PR 02-MAR-2001; 2001US-0272922P.
 PR 02-MAR-2001; 2001US-0273048P.
 PR 02-MAR-2001; 2001US-0273300P.
 PR 16-MAR-2001; 2001US-0276401P.
 PR 20-MAR-2001; 2001US-0277324P.

PR 20-MAR-2001; 2001US-0278660P.
 PR 30-MAR-2001; 2001US-0280039P.
 PR 30-MAR-2001; 2001US-0280234P.
 PR 02-APR-2001; 2001US-0280818P.
 PR 12-APR-2001; 2001US-0283443P.
 PR 23-APR-2001; 2001US-0285754P.
 PR 24-APR-2001; 2001US-0286096P.
 PR 03-MAY-2001; 2001US-0288353P.
 PR 17-MAY-2001; 2001US-0291703P.
 PR 31-MAY-2001; 2001US-0294834P.
 PR 20-JUN-2001; 2001US-0299695P.
 PR 21-JUN-2001; 2001US-0299845P.
 PR 05-JUL-2001; 2001US-0303242P.
 PR 13-AUG-2001; 2001US-0311981P.
 PR 16-AUG-2001; 2001US-0312858P.
 PR 17-AUG-2001; 2001US-0313280P.
 PR 29-AUG-2001; 2001US-0315614P.
 PR 17-SEP-2001; 2001US-0322818P.
 PR 25-FEB-2002; 2002US-00322818.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alsobrook JP, Anderson DW, Ballinger RA, Boldog FL, Burgess CE;
 PI Casman SJ, Ellerman KE, Gangolli EA, Gerlach VL, Gilbert JA;
 PI Gorman L, Guo X, Gusev VV, Kekuda R, Li L, Liu X, Malyankar UM;
 PI Miller CE, Millet I, Padigaru M, Patturajan M, Pena CEA, Peyman JA;
 PI Rastelli L, Shenoy SG, Shimmets RA, Smichson G, Spytek KA, Stone DJ;
 PI Taupier RJ, Tchernev VT, Vernet CAM, Zethusen BD;
 XX
 XX WPI; 2002-698672/75.
 DR N-PSDB; ADH48853.
 XX
 PT New NOVX polypeptides or polynucleotides, useful for preventing or
 PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
 PT obesity or cancer.
 XX
 BS Claim 1; Page 335; 923pp; English.
 XX
 CC The present invention relates to novel human NOVX proteins, where X is
 CC any number from 1 to 91 and their coding sequences. The proteins and
 CC coding sequences are useful for preventing or treating disorders or
 CC syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOV60 is
 CC a beta tectorin-like protein and its coding sequence maps to chromosome
 CC 10.
 XX
 SQ Sequence 300 AA;
 Query Match 87.7%; Score 1539.5; DB 5; Length 300;
 Best Local Similarity 89.4%; Pred. No. 4e-157;
 Matches 295; Conservative 2; Mismatches 2; Indels 31; Gaps 3;
 QY 1 MVTAFVLLAIFAEASAKSCAPNKADVLVFCYPTKIITIKIPECPYGEVHQLALGGLCY 60
 Db 1 MVTAFVLLAIFAEASAKSCAPNKADVLVFCYPTKIITIKIPECPYGEVHQLALGGLCY 60
 QY 61 NGVHEGGYQFVPIPLSPKNKSYCGTQSEYKPPYHFSYSHIVSNDTIVVKNQPNYSFS 120
 Db 61 NGVHEGGYQFVPIPLSPKNKSYCGTQSEYKPPYHFSYSHIVSNDTIVVKNQPNYSFS 120
 QY 121 CTYHSTYLVNQAAFDQVRVATVHVKNKSGMTFESQLSLNFY-TNAKFSIKKEAPFVLEA 179
 Db 121 CTYHSTYLVNQAAFDQ-----SYNFLPKNAKFSIKKEAPFVLEA 161
 QY 180 IGSDLFAGVEAKGLSIRFKVVLNSCWATPSADPMYPLQWLKNGCPTDETIVLVHNGRD 239
 Db 162 IGSDLFAGVEAKGLSIRFKVVLNSCWATPSADPMYPLQWLKNGCPTDETIVLVHNGRD 221
 QY 240 HRATFQFNAPRFONIPKLSKVLHCHETFCIDSEKLSKCPVTCDKRLRLDQTGGVLVVEL 299
 Db 222 HRATFQFNAPRFONIPKLSKVLHCHETFCIDSEKLSKCPVTCDKRLRLDQTGGVLVVEL 281
 QY 300 LRSRGFSSLYSFDVLHLLIMMLGICAVL 329
 ||| :|||||||

PS Example 48; SEQ ID NO 1060; 1880pp; English.

XX The invention relates to a novel isolated polypeptide (NOVX). A

CC polypeptide of the invention has cytostatic, immunomodulator,

CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and

CC antilipemic activity, and may have a use in gene therapy, and as a

CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising

CC any of the 303 fully defined nucleotide sequences given in the

CC specification. The polypeptide is useful in the manufacture of a

CC medicament for treating a syndrome associated with a human disease. The

CC polypeptide, polynucleotide and antibody are useful in diagnosing,

CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,

CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious

CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are

CC further used as hybridisation probes, in chromosome mapping, tissue

CC typing, preventive medicine, and pharmacogenomics. The present sequence

XX represents a NOVX polypeptide of the invention.

SY Sequence 276 AA;

Query Match 85.7%; Score 1505; DB 8; Length 276;

Best Local Similarity 100.0%; Pred. No. 1.9e-153;

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 KSCAPNKADVILVFCYKPTIITKIPECYPGWVHQLGGLCYNGVHGYYQFVIPDLS 77

DB 1 KSCAPNKADVILVFCYKPTIITKIPECYPGWVHQLGGLCYNGVHGYYQFVIPDLS 60

QY 78 PKNKSYCGTQSYKPIPIHYFHYSHIVSNDTIVTKQPNVNSFSCYTHSYLVNQAAFDQR 137

DB 61 PKNKSYCGTQSYKPIPIHYFHYSHIVSNDTIVTKQPNVNSFSCYTHSYLVNQAAFDQR 120

QY 138 VATVHVKGSMGTFFESQSLNFTYTNAKFSIKKEAPFVLEASIGSDLPAGVEAKGLSIRF 197

DB 121 VATVHVKGSMGTFFESQSLNFTYTNAKFSIKKEAPFVLEASIGSDLPAGVEAKGLSIRF 180

QY 198 KVVLSNCWATPSADFMYPQLQMLINKGCTDTETVLVHENGDRHRATFQNAFRFQNIPLK 257

DB 181 KVVLSNCWATPSADFMYPQLQMLINKGCTDTETVLVHENGDRHRATFQNAFRFQNIPLK 240

QY 258 SKVWLHCETFFICDSKLSCPVTCDDKRLRLRDQTGG 293

DB 241 SKVWLHCETFFICDSKLSCPVTCDDKRLRLRDQTGG 276

RESULT 9

ADH72166

ID ADH72166 standard; protein; 273 AA.

AC ADH72166;

XX 25-MAR-2004 (first entry)

DE Human protein of the invention NOV48e SEQ ID NO:1062.

XX human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

XX dyslipidaemia.

OS Homo sapiens.

XX WO2003102155-A2.

PN 11-DEC-2003.

PD 03-JUN-2003; 2003WO-US017430.

PF 03-JUN-2002; 2002US-0385120P.

XX 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

XX 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386646P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0389123P.

PR 14-JUN-2002; 2002US-0389118P.

PR 14-JUN-2002; 2002US-0389120P.

PR 14-JUN-2002; 2002US-0389144P.

PR 14-JUN-2002; 2002US-0389146P.

PR 17-JUN-2002; 2002US-0389729P.

PR 17-JUN-2002; 2002US-0389742P.

PR 18-JUN-2002; 2002US-0389844P.

PR 19-JUN-2002; 2002US-0390006P.

PR 19-JUN-2002; 2002US-0390209P.

PR 21-JUN-2002; 2002US-0390763P.

PR 17-JUL-2002; 2002US-0396706P.

PR 06-AUG-2002; 2002US-0401628P.

PR 09-AUG-2002; 2002US-0402156P.

PR 09-AUG-2002; 2002US-0402256P.

PR 09-AUG-2002; 2002US-0402389P.

PR 12-AUG-2002; 2002US-0402786P.

PR 12-AUG-2002; 2002US-0402816P.

PR 12-AUG-2002; 2002US-0402821P.

PR 12-AUG-2002; 2002US-0402832P.

PR 13-AUG-2002; 2002US-0403448P.

PR 13-AUG-2002; 2002US-0403531P.

PR 13-AUG-2002; 2002US-0403532P.

PR 13-AUG-2002; 2002US-0403563P.

PR 13-AUG-2002; 2002US-0406317P.

PR 15-AUG-2002; 2002US-0403617P.

PR 26-AUG-2002; 2002US-0406182P.

PR 26-AUG-2002; 2002US-0406355P.

PR 27-AUG-2002; 2002US-0406240P.

PR 12-SEP-2002; 2002US-0410084P.

PR 20-SEP-2002; 2002US-0412528P.

PR 23-SEP-2002; 2002US-0412731P.

PR 30-SEP-2002; 2002US-0414801P.

PR 30-SEP-2002; 2002US-0414839P.

PR 30-SEP-2002; 2002US-0414840P.

PR 30-SEP-2002; 2002US-0414954P.

PR 09-OCT-2002; 2002US-0417186P.

PR 23-OCT-2002; 2002US-0417406P.

PR 23-OCT-2002; 2002US-0420639P.

PR 28-OCT-2002; 2002US-0421156P.

PR 31-OCT-2002; 2002US-0422690P.

PR 01-NOV-2002; 2002US-0423130P.

PR 05-NOV-2002; 2002US-0423798P.

PR 05-NOV-2002; 2002US-0423798P.

PR 12-NOV-2002; 2002US-0425453P.

XX

PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 XX in gene mapping.
 PS Claim 27; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dithp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX
 SQ Sequence 640 AA;

Query Match 12.6%; Score 221; DB 8; Length 640;
 Best Local Similarity 25.7%; Pred. No. 3.8e-14;
 Matches 66; Conservative 49; Mismatches 102; Indels 40; Gaps 9;
 QY 76 LSPKNSYCGTQSEYKPPYHSHVSNDDTVIVKQPNVYFSCTYHSTYLVN-QAAF 134
 DB 381 VTPARDGCGTWT-RNETHATYNTLYLADEIIRDLNIRINACSYPLDMKVSILKTL 439
 QY 135 DQVATVHVKNKSGMGTTFESQLS- - - - -NFYTNKPSIKKEAPVLEASIGDLPAG 187
 DB 440 QPMVSALNIRVGTGTFVRLMFPQSTYQYQSSVTLSTEAFLYVGTMLDGGDL- - - 496
 QY 188 VEAKGLSIRFKVNLNSCWATPSADFMYPQLQINKGCP- -TDETVLVHNGRDRHATFQ 245
 DB 497 - - - - -SRFALLMTNCTATPSSNATDPLKYFIQDRCPHTRDSTIQVVENGESSQGRFS 549
 QY 246 FNAFRQNI PKLSKVWLHCETFCIDSEKLSCPVTCDKRLRLRDOTGGVL- - -VVEL- 299
 DB 550 VQMFREFAG- -NYDLVYLHCEVYLCDTNEKCKPTCSG- - - - -TRFRSGSVIDQSRVNLNLP 603
 QY 300 - - - - -SLRSRGSFL 309
 DB 604 ITRKGQVATVSRFSSL 620

RESULT 15
 ID ADO59935 standard; protein; 643 AA.
 XX
 AC ADO59935;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Bovine UMOD.
 XX
 KW human; bovine; murine; rat; uromodulin; UMOD; assay; detection; mutation;
 KW wild-type; diagnosis; renal disease; predisposition;
 KW medullary cystic kidney disease 2; MCKD2;
 KW familial juvenile gouty nephropathy; FUGN; renal failure; hyperuricaemia;
 KW gouty arthritis; enuresis.
 XX
 OS Bos taurus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 105 /note= "Position of missense mutation"
 FT Misc-difference 150 /note= "Position of missense mutation"
 FT Misc-difference 219 /note= "Position of missense mutation"

FT /note= "Position of missense mutation"
 XX WO2004038377-A2.
 XX
 PD 06-MAY-2004.

XX 23-OCT-2003; 2003WO-US033957.
 XX
 XX 23-OCT-2002; 2002US-0420768P.
 PR 02-DEC-2002; 2002US-0430318P.
 PR

XX (UYPI-) UNIV PITTSBURGH.
 PA (UYWA-) UNIV WAKE FOREST.
 PA

PI Hart TC, Hart PS, Gorry M, Bleyer AJ;
 XX

DR WPI; 2004-399985/37.
 DR GENBANK; S75958.

PT Diagnosing renal diseases or a predisposition to renal diseases, e.g.
 PT renal failure, hyperuricaemia, gouty arthritis or enuresis comprises
 PT assaying a genetic material to detect a mutation in at least a copy of
 PT the uromodulin genetic sequence.
 XX

PS Disclosure; Fig 6; 24pp; English.

XX This sequence represents bovine uromodulin (UMOD). The cDNA encoding this
 CC sequence was used in the method of the invention for diagnosing a disease
 CC or a predisposition to contract a disease. The method comprises obtaining
 CC genetic material from a test subject and assaying the genetic material to
 CC detect a mutation in at least a copy of the UMOD genetic sequence, where
 CC the presence of a UMOD mutation supports a diagnosis of a disease or a
 CC predisposition to contract a disease within the patient. The method also
 CC comprises comparing the sequence of the genetic material of the sequence
 CC of the wild-type UMOD gene and identifying any differences between the
 CC sequence of the genetic material and the wild-type UMOD gene. The method
 CC is useful for diagnosing renal diseases or a predisposition to renal
 CC diseases, e.g. medullary cystic kidney disease 2 (MCKD2), familial
 CC juvenile gouty nephropathy (FUGN), renal failure, hyperuricaemia, gouty
 CC arthritis or enuresis.

XX Sequence 643 AA;

Query Match 12.5%; Score 220; DB 8; Length 643;
 Best Local Similarity 25.8%; Pred. No. 4.9e-14;
 Matches 59; Conservative 46; Mismatches 98; Indels 26; Gaps 7;

QY 76 LSPKNSYCGTQSEYKPPYHSHVSNDDTVIVKQPNVYFSCTYHSTYLVN-QAAF 134
 DB 383 VTPARDGCGTWT-RNETHATYNTLYLADEIIRDLNIRINACSYPLDMKVSILKTL 441
 QY 135 DQVATVHVKNKSGMGTTFESQLS- - - - -NFYTNKPSIKKEAPVLEASIGDLPAG 187
 DB 442 QPMVSALNIRVGTGTFVRLMFPQSTYQYQSSVTLSTEAFLYVGTMLDGGDL- - - 498
 QY 188 VEAKGLSIRFKVNLNSCWATPSADFMYPQLQINKGCP- -TDETVLVHNGRDRHATFQ 245
 DB 499 - - - - -SRFVLLMTNCTATPSSNATDPLKYFIQDRCPHTRDSTIQVVENGESSQGRFS 551
 QY 246 FNAFRQNI PKLSKVWLHCETFCIDSEKLSCPVTCDK- - -RKRLRLDOT 291
 DB 552 VQMFREFAG- -NYDLVYLHCEVYLCDTNEKCKPTCPETFRSGSIIDOT 598

Search completed: July 20, 2005, 20:34:12
 Job time : 170 secs

THIS PAGE BLANK (USP 10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2005, 20:23:55 ; Search time 43 Seconds
(without alignments)
571.152 Million cell updates/sec

Title: US-10-687-268-35
Perfect score: 1756
Sequence: 1 MVTAKFVLIAFAEASAKS.....YFSVDLHMLHMLGICAVL 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/aaa/5A COMB.pep.*
- 2: /cgn2_6/prodata/1/aaa/5B COMB.pep.*
- 3: /cgn2_6/prodata/1/aaa/6A COMB.pep.*
- 4: /cgn2_6/prodata/1/aaa/6B COMB.pep.*
- 5: /cgn2_6/prodata/1/aaa/PCUS COMB.pep.*
- 6: /cgn2_6/prodata/1/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	227	12.9	614	4	US-09-949-016-8536
2	213	12.1	380	4	US-09-949-016-8534
3	213	12.1	380	4	US-09-949-016-8535
4	210	12.0	530	4	US-09-949-016-6085
5	210	12.0	554	4	US-09-949-016-8532
6	210	12.0	554	4	US-09-949-016-8533
7	210	12.0	554	4	US-09-949-016-9420
8	210	12.0	554	4	US-09-949-016-9421
9	209	11.9	384	1	US-08-350-435-2
10	209	11.9	384	5	PCT-US95-15916-2
11	172	9.8	607	4	US-09-907-794A-190
12	172	9.8	607	4	US-09-905-125A-190
13	172	9.8	607	4	US-09-902-775A-190
14	172	9.8	607	4	US-09-906-701A-190
15	172	9.8	607	4	US-09-903-603A-190
16	172	9.8	607	4	US-09-904-920A-190
17	172	9.8	607	4	US-09-909-064A-190
18	172	9.8	607	4	US-09-905-381A-190
19	172	9.8	607	4	US-09-906-618-190
20	155	8.8	254	4	US-09-808-701A-30
21	150.5	8.6	1785	3	US-09-341-587-3
22	149	8.5	545	4	US-10-067-422-11
23	148.5	8.5	566	2	US-08-484-993B-41
24	148.5	8.5	566	2	US-08-484-158B-41
25	148.5	8.5	566	2	US-08-484-596A-41
26	148.5	8.5	566	2	US-08-480-150A-41
27	148.5	8.5	566	3	US-08-458-731-41

28	148.5	8.5	566	3	US-08-149-223A-41	Sequence 41, Appl
29	148.5	8.5	713	2	US-08-484-993B-2	Sequence 2, Appl
30	148.5	8.5	713	2	US-08-484-158B-2	Sequence 2, Appl
31	148.5	8.5	713	2	US-08-484-596A-2	Sequence 2, Appl
32	148.5	8.5	713	2	US-08-480-150A-2	Sequence 2, Appl
33	148.5	8.5	713	3	US-08-458-731-2	Sequence 2, Appl
34	148.5	8.5	713	3	US-08-149-223A-2	Sequence 2, Appl
35	147.5	8.4	1290	1	US-08-470-350B-2	Sequence 6, Appl
36	144	8.2	713	1	US-08-453-472-6	Sequence 7, Appl
37	144	8.2	713	1	US-08-038-948-7	Sequence 8, Appl
38	144	8.2	713	1	US-08-038-948-8	Sequence 10, Appl
39	144	8.2	713	1	US-08-038-948-10	Sequence 6, Appl
40	144	8.2	713	1	US-08-453-952-6	Sequence 6, Appl
41	144	8.2	713	2	US-08-862-903-6	Sequence 14, Appl
42	142	8.1	716	2	US-08-484-993B-14	Sequence 14, Appl
43	142	8.1	716	2	US-08-484-158B-14	Sequence 14, Appl
44	142	8.1	716	2	US-08-484-596A-14	Sequence 14, Appl
45	142	8.1	716	2	US-08-480-150A-14	Sequence 14, Appl

RESULT 1
US-09-949-016-8536
; Sequence 8536, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8536
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-8536

Query Match 12.9%; Score 227; DB 4; Length 614;
Best Local Similarity 24.3%; Pred. No. 1.4e-15;
Matches 81; Conservative 55; Mismatches 127; Indels 70; Gaps 14;

QY	13	ABASAKCAPNKADVLVFCYPTIITIKIPCEPYGWEVHQLAGLCYN-GVHEGGYYQF 71
Db	296	ASVQVKACAGG-----YVYNLTAPPECHLAYCTDPSVSGTCRECSIDE----- 340
QY	72	VIPDLSPGN-KSYCG-----TOSEYKP-----PIYHPYSHIVNDTIVVKNQPVNYSF 119
Db	341	--DCKSNNGRWHCCQCKQDFNITAPEHKPGPRNETHATYSNTLYLADIIIRDINKINF 397
QY	120	SCYTHSTYLVN-OAADFQVATVHVKGSGMCTPESQSL-----NFTYNAKFSIKKEA 171
Db	398	ACSPFLDMKSLKTLQPMWSALNIRVGCTGMFTVRNALFOTPSYQYQSSVTLSTEA 457
QY	172	PFVLEASIEIGDLFAGVEAKGLSIRFKVNLNSCWATPSADPMYPLQWOLINKGCP--TDE 229
Db	458	FLYVGTMLDGGDL-----SRFALLMTCVATPSSNATDPLKYFIIDRCFHTRDS 507
QY	230	TVLVHENGDRIRATFQFNARFQNIPLKSLVHLHCETFI CDSEKLSCEVTCCKRLLRD 289
Db	508	TIQVENGESSQGRFSVQMFERFAG--NYDLVYLHCEVYLCDTMNEKCKPTCSG----TRF 561
QY	290	QTGGVL----VVEL-----SLSRSGFSSL 309

ALIGNMENTS

Db 265 WSVTSFVQASACRNIL-RNTHAIYKNTLSLVNDFFIIRDTILNINFCAYPLDMKVL 323
QY 131 QAAFDQVATVHVKNKSGMTFESQLS---NFYTNAKFSIKKEAPFVLEASIG--SDLF 185
Db 324 QAAQLPVSLSNVSDGNGEFIVRMALFDQNTN-----PYEGDAVELSVESVLY 374
QY 186 AG-VEAKGLSIRPKVVLNSCWATPSADFMYPQLWLNKGCPT--DETVLVHENGDRHRA 242
Db 375 VGAILEQDTSRFLNLCYATPTEDKADLVKYFIIRNSCSNQRDSTIHVEENGQSSS 434
QY 243 TQFNAPFRQNPKLKSKVWLHCETFCIDSEKLSCPVTCDK 282
Db 435 RFSVQMFAGHYDL--VFLHCEIHLCDLSLNEQCQPCSCR 472

RESULT 5

US-09-949-016-8532
; Sequence 8532, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8532
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8532

Query Match 12.0%; Score 210; DB 4; Length 554;
Best Local Similarity 26.8%; Pred. No. 8.4e-14;
Matches 75; Conservative 43; Mismatches 104; Indels 59; Gaps 11;
QY 32 CYPKTIITKIPCEPYGWEVHQLAGGLCYNGHGGYQFVIPDL----- 76
Db 246 CGPREIKVKVDC-----LLGGL-----GLGEEVIAYLRDPNCSSILOTEERN 288
QY 77 -----SPKNKSYCGTQSEYKPPHYFYSHIVSNDTTVIVKNQPVNVSFCTVHSTYLVN- 130
Db 289 WSVTSFVQASACRNIL-RNTHAIYKNTLSLVNDFFIIRDTILNINFCAYPLDMKVL 347
QY 131 QAAFDQVATVHVKNKSGMTFESQLS---NFYTNAKFSIKKEAPFVLEASIG--SDLF 185
Db 348 QAAQLPVSLSNVSDGNGEFIVRMALFDQNTN-----PYEGDAVELSVESVLY 398
QY 186 AG-VEAKGLSIRPKVVLNSCWATPSADFMYPQLWLNKGCPT--DETVLVHENGDRHRA 242
Db 399 VGAILEQDTSRFLNLCYATPTEDKADLVKYFIIRNSCSNQRDSTIHVEENGQSSS 458
QY 243 TQFNAPFRQNPKLKSKVWLHCETFCIDSEKLSCPVTCDK 282
Db 459 RFSVQMFAGHYDL--VFLHCEIHLCDLSLNEQCQPCSCR 496

RESULT 6

US-09-949-016-8533
; Sequence 8533, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8533
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8533

Query Match 12.0%; Score 210; DB 4; Length 554;
Best Local Similarity 26.8%; Pred. No. 8.4e-14;
Matches 75; Conservative 43; Mismatches 104; Indels 59; Gaps 11;
QY 32 CYPKTIITKIPCEPYGWEVHQLAGGLCYNGHGGYQFVIPDL----- 76
Db 246 CGPREIKVKVDC-----LLGGL-----GLGEEVIAYLRDPNCSSILOTEERN 288
QY 77 -----SPKNKSYCGTQSEYKPPHYFYSHIVSNDTTVIVKNQPVNVSFCTVHSTYLVN- 130
Db 289 WSVTSFVQASACRNIL-RNTHAIYKNTLSLVNDFFIIRDTILNINFCAYPLDMKVL 347
QY 131 QAAFDQVATVHVKNKSGMTFESQLS---NFYTNAKFSIKKEAPFVLEASIG--SDLF 185
Db 348 QAAQLPVSLSNVSDGNGEFIVRMALFDQNTN-----PYEGDAVELSVESVLY 398
QY 186 AG-VEAKGLSIRPKVVLNSCWATPSADFMYPQLWLNKGCPT--DETVLVHENGDRHRA 242
Db 399 VGAILEQDTSRFLNLCYATPTEDKADLVKYFIIRNSCSNQRDSTIHVEENGQSSS 458
QY 243 TQFNAPFRQNPKLKSKVWLHCETFCIDSEKLSCPVTCDK 282
Db 459 RFSVQMFAGHYDL--VFLHCEIHLCDLSLNEQCQPCSCR 496

RESULT 7

US-09-949-016-9420
; Sequence 9420, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9420
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9420

Query Match 12.0%; Score 210; DB 4; Length 554;
Best Local Similarity 26.8%; Pred. No. 8.4e-14;
Matches 75; Conservative 43; Mismatches 104; Indels 58; Gaps 11;
QY 32 CYPKTIITKIPCEPYGWEVHQLAGGLCYNGHGGYQFVIPDL----- 76

```

; Patent No. 5663315
; GENERAL INFORMATION:
; APPLICANT: Scheele, George
; APPLICANT: Fukuoka, Shin-Ichi
; TITLE OF INVENTION: GP2 and Diagnosis of
; TITLE OF INVENTION: Pancreatitis
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/350,435
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 06184/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-350-435-2

Query Match 11.9%; Score 209; DB 1; Length 384;
Best Local Similarity 24.8%; Pred. No. 6.2e-14;
Matches 88; Conservative 49; Mismatches 116; Indels 102; Gaps 15;

QY 1 MVTAKFVLLA---IFAEASA-----KSCAP----- 22
DB 1 MVSGLLWLVLCILTOASAVQVRDPDPSTVEKKEKACRPEECALNSTWGCRCRQ 60
QY 23 --NKADV---ILVFCYPTKIITKIPBCPYGWEVHQLAGLCYNGVHEGGYQFVLPDL 76
DB 61 DLNSDSVHSLQPLDCGPREIKVKVDC-----LQGL-----GFPEEVIAYL 103
QY 77 -----SPKNKSYCGTQSEYKPIYHYSHIVSNDTIVVKNQPVN 116
DB 104 RDPNCSSILOPEERNVWVTSFVQASACRNILE--RNOHTAIYKNTLSLVNDFIIRDTILN 162
QY 117 YSFCTVHSTVLVN-QAADFQVATVHVGSGMTFESQLS---NFTNAKFSIKKEAP 172
DB 163 INFOCAVPLDMKVSQALQPIVSSLVNVDGNGEIVRMALFQDQNTN-----P 213
QY 173 FVLEASIEG--SDLFAG-VEAKGLSIRFKVVLNSCWATPSADFMYPLOWQLINKGCPT-- 227
DB 214 YEGDAVELSVESVLYVGAILEQDTSRFLNLRNCYATPTEDKADLVKYFIIRNSCSNR 273
QY 228 DETVLVHNGRDHRAATQFNAPFRQNTPKLSKWLHCETFTICDSEKLSCPVTCCK 282
DB 274 DSTIIVHENGQSSRSFVQMFPMFAGHYDL--VFLHCEIHLCDLSLNEQCQPSCSR 326

RESULT 10
PCT-US95-15916-2
; Sequence 2, Application PC/TUS9515916
; GENERAL INFORMATION:
; APPLICANT: ALPHAGEN, INC.
; TITLE OF INVENTION: Diagnosis of Pancreatitis

```

```

; Patent No. 5663315
; GENERAL INFORMATION:
; APPLICANT: Scheele, George
; APPLICANT: Fukuoka, Shin-Ichi
; TITLE OF INVENTION: GP2 and Diagnosis of
; TITLE OF INVENTION: Pancreatitis
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/350,435
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 06184/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-350-435-2

Query Match 12.0%; Score 210; DB 4; Length 554;
Best Local Similarity 26.8%; Pred. No. 8.4e-14;
Matches 75; Conservative 43; Mismatches 104; Indels 58; Gaps 11;

QY 32 CYPKTIITKIPBCPYGWEVHQLAGLCYNGVHEGGYQFVLPDL----- 76
DB 246 CGPREIKVKVDC-----LLGL-----GLGEEVIAYLRDPNCSILQTEERN 288
QY 77 -----SPKNKSYCGTQSEYKPIYHYSHIVSNDTIVVKNQPVNVSFCTVHSTVLVN- 130
DB 289 WVSFVTSFVQASACRNILE-RNQHTAIYKNTLSLVNDFIIRDTILNINFCAYPLDMKVS 347
QY 131 QAAPQVATVHVGSGMTFESQLS---NFTNAKFSIKKEAPFVLEASIEG--SDLF 185
DB 348 QAALQPIVSSLVNVDGNGEIVRMALFQDQNTN-----PYEGDAVELSVESVLY 398
QY 186 AG-VEAKGLSIRFKVVLNSCWATPSADFMYPLOWQLINKGCPT--DETVLVHNGRDHRA 242
DB 399 VGAILEQDTSRFLNLRNCYATPTEDKADLVKYFIIRNSCSNRQDSTIIVHENGQSS 458
QY 243 TQFNAFRQNTPKLSKWLHCETFTICDSEKLSCPVTCCK 282
DB 459 RFSVQMFPMFAGHYDL--VFLHCEIHLCDLSLNEQCQPSCSR 496

RESULT 8
US-09-949-016-9421
; Sequence 9421, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9421
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-9421

```

```

; Sequence 2, Application US/08350435

```


Db 264 SIYAENINTSLTSSDRMRVLSKSYL-----AFNS--NG 298
Qy 67 GYQFVLPDLSPKNSY-----CGTQSEYKPPYVHFYSHIV-----SNDTTVIIVKNQ 113
Db 299 NNQLQKDPCTCPKLSNVVFSVPLNGCGTIRKVEDQSI-TYTNITFSASSTSEVITRQK 357
Qy 114 PWNYSFCT--YHSTYLVNQAFQDQVATVHVKNQSMGTFFESQSLNFVYTNKFSIKKEA 171
Db 358 QLOIIVKCEMGNSTVEIYITDDVIQS-----QNALGKYNTSMAL-FESNSFEKTILES 412
Qy 172 PFVLEASIEGDLFAGVEAKGLSIRFKVVLNSCWATPSADFMYPQLQWLNKNGCPTDET 231
Db 413 PYVY---DLNQTFLVQVSLHTSDPNLVFLDTCRASPTSDPASP-TYDLIKSGCSRDTC 468
Qy 232 LVHENGDRHRATFQNAFRFQNIPLKSKVWLHCETFCIDS--EKUSCPVTCDKRRK 285
Db 469 KYVPL-FGHYGRFQFNAKFF--LRSMSSVYLQCKVLICDSSDHQSRQNCQGVRSRK 521

RESULT 12
US-09-905-125A-190
; Sequence 190, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Goddard, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 190
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-125A-190

Query Match 9.8%; Score 172; DB 4; Length 607;
Best Local Similarity 23.3%; Pred. No. 1.3e-09;
Matches 69; Conservative 54; Mismatches 115; Indels 58; Gaps 14;

Qy 10 AIPAE---ASAKSCAPNKADVILVFCYPKTIITKIPECYPGWVHQLALGCLCYNGVHEG 66
Db 264 SIYAENINTSLTSSDRMRVLSKSYL-----AFNS--NG 298
Qy 67 GYQFVLPDLSPKNSY-----CGTQSEYKPPYVHFYSHIV-----SNDTTVIIVKNQ 113
Db 299 NNQLQKDPCTCPKLSNVVFSVPLNGCGTIRKVEDQSI-TYTNITFSASSTSEVITRQK 357
Qy 114 PWNYSFCT--YHSTYLVNQAFQDQVATVHVKNQSMGTFFESQSLNFVYTNKFSIKKEA 171
Db 358 QLOIIVKCEMGNSTVEIYITDDVIQS-----QNALGKYNTSMAL-FESNSFEKTILES 412
Qy 172 PFVLEASIEGDLFAGVEAKGLSIRFKVVLNSCWATPSADFMYPQLQWLNKNGCPTDET 231
Db 413 PYVY---DLNQTFLVQVSLHTSDPNLVFLDTCRASPTSDPASP-TYDLIKSGCSRDTC 468
Qy 232 LVHENGDRHRATFQNAFRFQNIPLKSKVWLHCETFCIDS--EKUSCPVTCDKRRK 285
Db 469 KYVPL-FGHYGRFQFNAKFF--LRSMSSVYLQCKVLICDSSDHQSRQNCQGVRSRK 521

RESULT 13
US-09-902-775A-190
; Sequence 190, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann


```

; LENGTH: 607
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-906-700-190

Query Match
Best Local Similarity 9.8%; Score 172; DB 4; Length 607;
Matches 69; Conservative 54; Mismatches 115; Indels 58; Gaps 14;

QY 10 AIFAE---ASAKCAPNKADVILVFCYKPTIITKIPECPYGVHQLALGCLYGVHHEG 66
Db 264 SIYAENINTSLTSSDRMRVLSKSYLE-----AFNS--NG 298

QY 67 GYQFVIPDLSPKNSY-----CCTQSEYKPPYHYFYSHV-----SNDTTVIVKQ 113
Db 299 NNQLQKPTCRPKLSNVVFSVPLNGCGTIRKVEDQSI--TYNIIITFSASSTSEVITRQK 357

QY 114 PNYISFSC--YHSTYLVNQAAFDQVATVHVKNMGSMGTFFESQLSLNFTYNAKFSIKKEA 171
Db 358 QLOIIVKCEMGNSTVEIYITDDVIQS-----QNALGKYNTSMAL-FESNSFEKTILES 412

QY 172 PFVLEASIGSLFAGVEAKGLSIRFKVLNSCWATPSADPMYPLQWLINKGCPDTDTV 231
Db 413 PYV---DLNQTLFVQVSLHTSDPNLVFLDTCRASPTSDPASP--TYDLIKSGGSRDTC 468

QY 232 LVHENGDRHRATFOFNAPRFQNIPLKSKVWLHCETFFICDS--EKLSCPVTCDKRRK 285
Db 469 KYVPL-FGHYGRFQFNAKFP--LRSMSSVYLQCKVLICDSDHQRNCGVSRSK 521

```

RESULT 15

```

US-09-903-603A-190
; Sequence 190, Application US/09903603A
; Patent No. 6767995
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, David
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: GNE.1618P2C12
; CURRENT APPLICATION NUMBER: US/09/903,603A
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28

```

```

; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 190
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-903-603A-190

Query Match
Best Local Similarity 9.8%; Score 172; DB 4; Length 607;
Matches 69; Conservative 54; Mismatches 115; Indels 58; Gaps 14;

QY 10 AIFAE---ASAKCAPNKADVILVFCYKPTIITKIPECPYGVHQLALGCLYGVHHEG 66
Db 264 SIYAENINTSLTSSDRMRVLSKSYLE-----AFNS--NG 298

QY 67 GYQFVIPDLSPKNSY-----CCTQSEYKPPYHYFYSHV-----SNDTTVIVKQ 113
Db 299 NNQLQKPTCRPKLSNVVFSVPLNGCGTIRKVEDQSI--TYNIIITFSASSTSEVITRQK 357

QY 114 PNYISFSC--YHSTYLVNQAAFDQVATVHVKNMGSMGTFFESQLSLNFTYNAKFSIKKEA 171
Db 358 QLOIIVKCEMGNSTVEIYITDDVIQS-----QNALGKYNTSMAL-FESNSFEKTILES 412

QY 172 PFVLEASIGSLFAGVEAKGLSIRFKVLNSCWATPSADPMYPLQWLINKGCPDTDTV 231
Db 413 PYV---DLNQTLFVQVSLHTSDPNLVFLDTCRASPTSDPASP--TYDLIKSGGSRDTC 468

QY 232 LVHENGDRHRATFOFNAPRFQNIPLKSKVWLHCETFFICDS--EKLSCPVTCDKRRK 285
Db 469 KYVPL-FGHYGRFQFNAKFP--LRSMSSVYLQCKVLICDSDHQRNCGVSRSK 521

```

Search completed: July 20, 2005, 20:37:54
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2005, 20:31:30 ; Search time 157 Seconds
(without alignments)
813.464 Million cell updates/sec

Title: US-10-687-268-35
Perfect score: 1756
Sequence: 1 MVTRAFVLLAFAEASAKS.....YFSDVHLHMLGICAVL 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1736639 seqs, 388189149 residues

Total number of hits satisfying chosen parameters: 1736639

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1756	100.0	329 15	US-10-312-088-35
2	1756	100.0	329 15	Sequence 35, Appl
3	1756	100.0	329 18	US-10-085-198-142
4	1751	99.7	329 15	Sequence 35, Appl
5	1665	94.8	329 18	US-10-085-198-140
6	1539.5	87.7	300 15	Sequence 140, Appl
7	219	12.5	577 15	Sequence 46, Appl
8	219	12.5	673 15	Sequence 138, Appl
9	213	12.1	406 9	US-10-104-047-2774
10	210	12.0	530 9	Sequence 3485, App
11	210	12.0	534 17	Sequence 2774, App
				Sequence 726, App
				Sequence 17, Appl
				Sequence 52, Appl

12	174	9.9	620	15	US-10-187-975-110	Sequence 110, App
13	173.5	9.9	1011	15	US-10-161-493-136	Sequence 136, App
14	172	9.8	530	15	US-10-330-051A-70	Sequence 70, Appl
15	172	9.8	585	9	US-09-864-711-14	Sequence 14, Appl
16	172	9.8	607	9	US-09-909-320-190	Sequence 190, App
17	172	9.8	607	9	US-09-909-088B-190	Sequence 190, App
18	172	9.8	607	9	US-09-905-291A-190	Sequence 190, App
19	172	9.8	607	9	US-09-902-853-190	Sequence 190, App
20	172	9.8	607	9	US-09-907-824-190	Sequence 190, App
21	172	9.8	607	9	US-09-907-841-190	Sequence 190, App
22	172	9.8	607	10	US-09-904-011-190	Sequence 190, App
23	172	9.8	607	10	US-09-903-640-190	Sequence 190, App
24	172	9.8	607	10	US-09-908-093-190	Sequence 190, App
25	172	9.8	607	10	US-09-906-742-190	Sequence 190, App
26	172	9.8	607	10	US-09-906-838-190	Sequence 190, App
27	172	9.8	607	10	US-09-907-613-190	Sequence 190, App
28	172	9.8	607	10	US-09-907-942-190	Sequence 190, App
29	172	9.8	607	10	US-09-904-859-190	Sequence 190, App
30	172	9.8	607	10	US-09-909-204-190	Sequence 190, App
31	172	9.8	607	10	US-09-904-820-190	Sequence 190, App
32	172	9.8	607	10	US-09-904-786-190	Sequence 190, App
33	172	9.8	607	10	US-09-906-646-190	Sequence 190, App
34	172	9.8	607	10	US-09-906-700-190	Sequence 190, App
35	172	9.8	607	10	US-09-903-786-190	Sequence 190, App
36	172	9.8	607	10	US-09-902-903-190	Sequence 190, App
37	172	9.8	607	10	US-09-903-749A-190	Sequence 190, App
38	172	9.8	607	10	US-09-904-119-190	Sequence 190, App
39	172	9.8	607	10	US-09-904-956-190	Sequence 190, App
40	172	9.8	607	10	US-09-902-736-190	Sequence 190, App
41	172	9.8	607	10	US-09-907-794-190	Sequence 190, App
42	172	9.8	607	10	US-09-903-943-190	Sequence 190, App
43	172	9.8	607	10	US-09-904-462-190	Sequence 190, App
44	172	9.8	607	10	US-09-907-925-190	Sequence 190, App
45	172	9.8	607	10	US-09-902-692-190	Sequence 190, App

ALIGNMENTS

RESULT 1
US-10-312-088-35
; Sequence 35, Application US/10312088
; Publication No. US20030219862A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Cogswell, John P.
; APPLICANT: Kabnic, Karen S.
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Martensen, Shelby A.
; APPLICANT: Murdock, Paul R.
; APPLICANT: Smith, Randall F.
; APPLICANT: Strum, Jay C.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Rizni, Safia K.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50029
; CURRENT APPLICATION NUMBER: US/10/312,088
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-312-088-35

Query Match 100.0%; Score 1756; DB 15; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.9e-175;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTAFVLLAIFAASAKSCAPNKADVLVFCYPTKIITKIPECPYGWEVHQLALGGLCY 60
DB 1 MVTAFVLLAIFAASAKSCAPNKADVLVFCYPTKIITKIPECPYGWEVHQLALGGLCY 60

QY 61 NGVHEGGYYQFVLPDLSPKNKSYCGTQSEYKPPYIHFYSHIVSNDTIVVKNQPNVYSFS 120
DB 61 NGVHEGGYYQFVLPDLSPKNKSYCGTQSEYKPPYIHFYSHIVSNDTIVVKNQPNVYSFS 120

QY 121 CTYHSTYLVNQAAFDQVRVATVHVKNKSGMGTTFESQSLNFTYNAKFSIKKEAPFVLEASEI 180
DB 121 CTYHSTYLVNQAAFDQVRVATVHVKNKSGMGTTFESQSLNFTYNAKFSIKKEAPFVLEASEI 180

QY 181 GSDLPAGVEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCPTDETIVLVHENGDRH 240
DB 181 GSDLPAGVEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCPTDETIVLVHENGDRH 240

QY 241 RATFOFNAFRFQNIPLSKVWLHCETFCIDSEKLSCPVTCDKRKLRLRDQTGGVVLVVELS 300
DB 241 RATFOFNAFRFQNIPLSKVWLHCETFCIDSEKLSCPVTCDKRKLRLRDQTGGVVLVVELS 300

QY 301 LRSRGFSSLYSFSVDLHLLIMMLGICAVL 329
DB 301 LRSRGFSSLYSFSVDLHLLIMMLGICAVL 329

RESULT 2

US-10-085-198-142
; Sequence 142, Application US/10085198
; Publication No. US2004000907A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-279
; CURRENT APPLICATION NUMBER: US/10/085,198
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/271,646
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/276,401
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,981
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,858
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/271,840
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/277,324
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/286,096
; PRIOR FILING DATE: 2001-04-21
; PRIOR APPLICATION NUMBER: 60/299,695
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/315,614
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/272,405
; PRIOR FILING DATE: 2001-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 653
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 142
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-085-198-142

Query Match 100.0%; Score 1756; DB 15; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.9e-175;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTAFVLLAIFAASAKSCAPNKADVLVFCYPTKIITKIPECPYGWEVHQLALGGLCY 60

DB 1 MVTAFVLLAIFAASAKSCAPNKADVLVFCYPTKIITKIPECPYGWEVHQLALGGLCY 60

QY 61 NGVHEGGYYQFVLPDLSPKNKSYCGTQSEYKPPYIHFYSHIVSNDTIVVKNQPNVYSFS 120

DB 61 NGVHEGGYYQFVLPDLSPKNKSYCGTQSEYKPPYIHFYSHIVSNDTIVVKNQPNVYSFS 120

QY 121 CTYHSTYLVNQAAFDQVRVATVHVKNKSGMGTTFESQSLNFTYNAKFSIKKEAPFVLEASEI 180

DB 121 CTYHSTYLVNQAAFDQVRVATVHVKNKSGMGTTFESQSLNFTYNAKFSIKKEAPFVLEASEI 180

QY 181 GSDLPAGVEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCPTDETIVLVHENGDRH 240

DB 181 GSDLPAGVEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCPTDETIVLVHENGDRH 240

QY 241 RATFOFNAFRFQNIPLSKVWLHCETFCIDSEKLSCPVTCDKRKLRLRDQTGGVVLVVELS 300

DB 241 RATFOFNAFRFQNIPLSKVWLHCETFCIDSEKLSCPVTCDKRKLRLRDQTGGVVLVVELS 300

QY 301 LRSRGFSSLYSFSVDLHLLIMMLGICAVL 329

DB 301 LRSRGFSSLYSFSVDLHLLIMMLGICAVL 329

RESULT 3

US-10-687-268-35
; Sequence 35, Application US/10687268
; Publication No. US20050137129A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Lee, Judithann M.
; APPLICANT: Smith, Randall F.
; APPLICANT: White, John R.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50029-1
; CURRENT APPLICATION NUMBER: US/10/687,268
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 10/312,088
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-687-268-35

Query Match 100.0%; Score 1756; DB 18; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.9e-175;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTAFVLLAIFAASAKSCAPNKADVLVFCYPTKIITKIPECPYGWEVHQLALGGLCY 60

DB 1 MVTAFVLLAIFAASAKSCAPNKADVLVFCYPTKIITKIPECPYGWEVHQLALGGLCY 60

QY 61 NGVHEGGYYQFVLPDLSPKNKSYCGTQSEYKPPYIHFYSHIVSNDTIVVKNQPNVYSFS 120

DB 61 NGVHEGGYYQFVLPDLSPKNKSYCGTQSEYKPPYIHFYSHIVSNDTIVVKNQPNVYSFS 120

QY 121 CTYHSTYLVNQAAFDQVRVATVHVKNKSGMGTTFESQSLNFTYNAKFSIKKEAPFVLEASEI 180

DB 121 CTYHSTYLVNQAAFDQVRVATVHVKNKSGMGTTFESQSLNFTYNAKFSIKKEAPFVLEASEI 180

QY 181 GSDLPAGVEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCPTDETIVLVHENGDRH 240

DB 181 GSDLPAGVEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCPTDETIVLVHENGDRH 240

QY 241 RATQFNARFQNIPLKSKVWLHCETFCDSKLSCPVTCCKRKLRLRDQGTGVLVVELS 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 RATQFNARFQNIPLKSKVWLHCETFCDSKLSCPVTCCKRKLRLRDQGTGVLVVELS 300
QY 301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329

RESULT 4
US-10-085-198-140
; Sequence 140, Application US/10085198
; Publication No. US2004000907A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; CURRENT APPLICATION NUMBER: US/10/085,198
; FILE REFERENCE: 21402-279
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/271,646
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/276,401
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,981
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,858
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/271,840
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/277,324
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/286,096
; PRIOR FILING DATE: 2001-04-21
; PRIOR APPLICATION NUMBER: 60/299,695
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/315,614
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/272,405
; PRIOR FILING DATE: 2001-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 653
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 140
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-085-198-140

Query Match 99.7%; Score 1751; DB 15; Length 329;
Best Local Similarity 99.7%; Pred. No. 6.3e-175;
Matches 328; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MVTAFVLLAIPAEASAKSCAPNKADVLVFCYPKTIITKIPECYGVHQLALGGLCY 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 MVTAFVLLAIPAEASAKSCAPNKADVLVFCYPKTIITKIPECYGVHQLALGGLCY 60
QY 61 NGVHEGGYQFVLPDLSPKNSYCGTQSEYKPPYHFSYHIVSNDDTVIVKQPNYSFS 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 NGVHEGGYQFVLPDLSPKNSYCGTQSEYKPPYHFSYHIVSNDDTVIVKQPNYSFS 120
QY 121 CTYHSTYLVNQAFQDVATVHVKNKSGMTFESQSLNFTYNKFSIKKEAPFVLESEI 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 CTYHSTYLVNQAFQDVATVHVKNKSGMTFESQSLNFTYNKFSIKKEAPFVLESEI 180
QY 181 GSDLFAGVEAKGLSIRFKVNLSCWATPSADPMYPLQMLINKGCTPTDVLVHENGDRH 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 GSDLFAGVEAKGLSIRFKVNLSCWATPSADPMYPLQMLINKGCTPTDVLVHENGDRH 240
QY 241 RATQFNARFQNIPLKSKVWLHCETFCDSKLSCPVTCCKRKLRLRDQGTGVLVVELS 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 RATQFNARFQNIPLKSKVWLHCETFCDSKLSCPVTCCKRKLRLRDQGTGVLVVELS 300
QY 301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329

Db 301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329
|||
RESULT 5
US-10-687-268-46
; Sequence 46, Application US/10687268
; Publication No. US20050137129A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Lee, Judithann M.
; APPLICANT: Smith, Randall F.
; APPLICANT: White, John R.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50029-1
; CURRENT APPLICATION NUMBER: US/10/687,268
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 10/312,088
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-687-268-46

Query Match 94.8%; Score 1665; DB 18; Length 329;
Best Local Similarity 94.2%; Pred. No. 6.9e-166;
Matches 310; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
QY 1 MVTAFVLLAIPAEASAKSCAPNKADVLVFCYPKTIITKIPECYGVHQLALGGLCY 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 MVTAFVLLAIPAEASAKSCAPNKADVLVFCYPKTIITKIPECYGVHQLALGGLCY 60
QY 61 NGVHEGGYQFVLPDLSPKNSYCGTQSEYKPPYHFSYHIVSNDDTVIVKQPNYSFS 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 NGVHEGGYQFVLPDLSPKNSYCGTQSEYKPPYHFSYHIVSNDDTVIVKQPNYSFS 120
QY 121 CTYHSTYLVNQAFQDVATVHVKNKSGMTFESQSLNFTYNKFSIKKEAPFVLESEI 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 CTYHSTYLVNQAFQDVATVHVKNKSGMTFESQSLNFTYNKFSIKKEAPFVLESEI 180
QY 181 GSDLFAGVEAKGLSIRFKVNLSCWATPSADPMYPLQMLINKGCTPTDVLVHENGDRH 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 GSDLFAGVEAKGLSIRFKVNLSCWATPSADPMYPLQMLINKGCTPTDVLVHENGDRH 240
QY 241 RATQFNARFQNIPLKSKVWLHCETFCDSKLSCPVTCCKRKLRLRDQGTGVLVVELS 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 RATQFNARFQNIPLKSKVWLHCETFCDSKLSCPVTCCKRKLRLRDQGTGVLVVELS 300
QY 301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329

RESULT 6
US-10-085-198-138
; Sequence 138, Application US/10085198
; Publication No. US2004000907A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-279
; CURRENT APPLICATION NUMBER: US/10/085,198
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/271,646

```

; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/276,401
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,981
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,858
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/271,840
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/277,324
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/286,096
; PRIOR FILING DATE: 2001-04-21
; PRIOR APPLICATION NUMBER: 60/299,695
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/315,614
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/272,405
; PRIOR FILING DATE: 2001-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 653
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 138
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-085-198-138

Query Match      87.7%; Score 1539.5; DB 15; Length 300;
Best Local Similarity 89.4%; Pred. No. 9.2e-153;
Matches 295; Conservative 2; Mismatches 2; Indels 31; Gaps 3;

QY 1 MVTKAFVLLAFAEASAKSCAPNKADVLVFCYKPTIITKIPECPYGEVHQLAGGLCY 60
DB 1 MVTKAFVLLAFAEASAKSCAPNKADVLVFCYKPTIITKIPECPYGEVHQLAGGLCY 60
QY 61 NGVHEGGYQFVPLDSPKNKSYCTQSEYKPPPIYHFYSHIVNSDNTTVIVKQPNVYSFS 120
DB 61 NGVHEGGYQFVPLDSPKNKSYCTQSEYKPPPIYHFYSHIVNSDNTTVIVKQPNVYSFS 120
QY 121 CYYHSTYLVNQAFQFQVATVHVKNKSGMTFESQSLNFI--TNAKFSIKKEAPFVLEAS 179
DB 121 CYYHSTYLVNQAFQFQVATVHVKNKSGMTFESQSLNFI--TNAKFSIKKEAPFVLEAS 161
QY 180 IGSDFAGVEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCPDDETTLVHENGDR 239
DB 162 IGSDFAGVEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCPDDETTLVHENGDR 221
QY 240 HRATFQNAFRQNPILSKVWLHCETFCIDSEKLSCPVTCCKRKLRLDQGTGGVLVVEL 299
DB 222 HRATFQNAFRQNPILSKVWLHCETFCIDSEKLSCPVTCCKRKLRLDQGTGGVLVVEL 281
QY 300 SLRSRGFSSLSYFSFVLLHLLIMLIGICAVL 329
DB 282 SLR-----NVLLHLLIMLIGICAVL 300

RESULT 7
US-10-108-260A-3485
; Sequence 3485, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3485
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-108-260A-3485

```

Query Match      12.5%; Score 219; DB 15; Length 577;
Best Local Similarity 25.7%; Pred. No. 1.3e-13;
Matches 66; Conservative 49; Mismatches 102; Indels 40; Gaps 9;

QY 76 LSPKNKSYCGTQSEYKPPPIYHFYSHIVNSDNTTVIVKQPNVYSFSCYHSTYLVN-QAAF 134
DB 318 VTPARDGPGCTVLT-RNETHATYNTLYLADEIIIRDNLINIKINFACSYPLDMKVS LKAL 376
QY 135 DQVATVHVKNKSGMTFESQSL-----NFYTNAKFSIKKEAPFVLEASIGSDLFAG 187
DB 377 QPMVSALNIRVGGTGMFTVRMALTQPSYTPQYQSSVTLSTEAFVVGTMLDGGDL--- 433
QY 188 VEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCP--TDETTLVHENGDRHRAF 245
DB 434 -----SRFALLMNTCYATPSSNATDPLKYFIIDRCPTHTRDSTIQVVENGESSQGRFS 486
QY 246 FNAFRQNPILSKVWLHCETFCIDSEKLSCPVTCCKRKLRLDQGTGGVL-----VVEL-- 299
DB 487 VQMFRRFAG--NYDLVYLHCEVYLCDTNEKCKPTCSG----TRFRSGSVIDQSRVNLGP 540
QY 300 -----SLRSRGFSSL 309
DB 541 ITRKGQVATVSRAFSSL 557

```

RESULT 8

```

US-10-104-047-2774
; Sequence 2774, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2774
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2774

```

```

Query Match      12.5%; Score 219; DB 15; Length 673;
Best Local Similarity 25.7%; Pred. No. 1.6e-13;
Matches 66; Conservative 49; Mismatches 102; Indels 40; Gaps 9;

QY 76 LSPKNKSYCGTQSEYKPPPIYHFYSHIVNSDNTTVIVKQPNVYSFSCYHSTYLVN-QAAF 134
DB 414 VTPARDGPGCTVLT-RNETHATYNTLYLADEIIIRDNLINIKINFACSYPLDMKVS LKAL 472
QY 135 DQVATVHVKNKSGMTFESQSL-----NFYTNAKFSIKKEAPFVLEASIGSDLFAG 187
DB 473 QPMVSALNIRVGGTGMFTVRMALTQPSYTPQYQSSVTLSTEAFVVGTMLDGGDL--- 529
QY 188 VEAKGLSIRFKVNLNSCWATPSADFMYPLOWQLINKGCP--TDETTLVHENGDRHRAF 245
DB 530 -----SRFALLMNTCYATPSSNATDPLKYFIIDRCPTHTRDSTIQVVENGESSQGRFS 582
QY 246 FNAFRQNPILSKVWLHCETFCIDSEKLSCPVTCCKRKLRLDQGTGGVL-----VVEL-- 299
DB 583 VQMFRRFAG--NYDLVYLHCEVYLCDTNEKCKPTCSG----TRFRSGSVIDQSRVNLGP 636
QY 300 -----SLRSRGFSSL 309
DB 637 ITRKGQVATVSRAFSSL 653

RESULT 9

```

```

US-09-925-297-726
; Sequence 726, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 726
; LENGTH: 406
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (160)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-726

Query Match
Best Local Similarity 12.1%; Score 213; DB 9; Length 406;
Matches 89; Conservative 48; Mismatches 116; Indels 98; Gaps 15;

QY 1 MYTKAFVLLA-----KSCAP-----NK 24
DB 27 MVGSLGLWALSCILTAQAVQDPSTVEDKCEKACRPEEBCEALNLTWGCFCRQDLNS 86
QY 25 ADV-----ILVFCYKTIITIKIEPCYGVHQLGGLCYNGVHEGGYQFVLPDL-----76
DB 87 SDVHSLQPLDQCPREIKVKDKC-----LLGGL-----GLGEVIAYLRDPN 129
QY 77 -----SPKNKSYCGTQSEYKPIYHYFHYSHIVSNDTIVTKNQPNVNSPS 120
DB 130 CSSILOTEERNVSVTPVQASACRNILE-RNQTHAIYKNTLSLVNDFIIRTNINQ 188
QY 121 CTHSTYLVN-QAADFQVRVATVHVKNVSGMTFESQLSL---NFYTNKFSIKKEAPFVLE 176
DB 189 CAYPLDMKVSQAALQPIVSSLSNVSDGNGEIVRMALFQDQNYTN-----PYEGD 239
QY 177 ASEIG--SDLFAG-VEAKGLSRFKVLSNWCWATPSADPMYPLQWLKNGCPT--DET 231
DB 240 AVESVSVLYVGAILEQDTSRFLNLCVATPTEDKADLVKFIIRNSCSNQDSTI 299
QY 232 LVHENGDRHRTFQFNAPRFQNIPLSKVWLHCETFCIDSEKLSCPVTCDK 282
DB 300 HVENGGQSSSRFSVQMFAGHYDL--VFLHCEIHLCDLSNEQCPSCSR 348

RESULT 10
US-09-872-153-17
; Sequence 17, Application US/09872153
; Patent No. US20020082207A1
; GENERAL INFORMATION:
; APPLICANT: First, Shannon K.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, David C.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.531
; CURRENT APPLICATION NUMBER: US/09/872,153
; CURRENT FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 530
; TYPE: PR

Query Match
Best Local Similarity 12.0%; Score 210; DB 17; Length 534;
Matches 75; Conservative 43; Mismatches 104; Indels 58; Gaps 11;

QY 32 CYPKTIITIKIEPCYGVHQLGGLCYNGVHEGGYQFVLPDL-----76
DB 226 CGPREIKVKDKC-----LLGGL-----GLGEVIAYLRDPNCSILQTEERN 268
QY 77 -----SPKNKSYCGTQSEYKPIYHYFHYSHIVSNDTIVTKNQPNVNSFCTYHSTYLVN- 130
DB 269 WVSVTSPVQASACRNILE-RNQTHAIYKNTLSLVNDFIIRTNINQFQAYPLDMKVS 327
QY 131 QAADFQVRVATVHVKNVSGMTFESQLSL---NFYTNKFSIKKEAPFVLEASIG--SDLF 185
DB 328 QAALQPIVSSLSNVSDGNGEIVRMALFQDQNYTN-----PYEGDAVELSVESVLY 378
QY 186 AG-VEAKGLSRFKVLSNWCWATPSADPMYPLQWLKNGCPT--DET 242
DB 379 VGAILEQDTSRFLNLCVATPTEDKADLVKFIIRNSCSNQDSTI 438
QY 243 TFQFNAPRFQNIPLSKVWLHCETFCIDSEKLSCPVTCDK 282
DB 439 RFSVQMFAGHYDL--VFLHCEIHLCDLSNEQCPSCSR 476

```

```

; ORGANISM: Homo sapien
US-09-872-153-17

Query Match
Best Local Similarity 12.0%; Score 210; DB 9; Length 530;
Matches 75; Conservative 43; Mismatches 104; Indels 58; Gaps 11;

QY 32 CYPKTIITIKIEPCYGVHQLGGLCYNGVHEGGYQFVLPDL-----76
DB 222 CGPREIKVKDKC-----LLGGL-----GLGEVIAYLRDPNCSILQTEERN 264
QY 77 -----SPKNKSYCGTQSEYKPIYHYFHYSHIVSNDTIVTKNQPNVNSFCTYHSTYLVN- 130
DB 265 WVSVTSPVQASACRNILE-RNQTHAIYKNTLSLVNDFIIRTNINQFQAYPLDMKVS 323
QY 131 QAADFQVRVATVHVKNVSGMTFESQLSL---NFYTNKFSIKKEAPFVLEASIG--SDLF 185
DB 324 QAALQPIVSSLSNVSDGNGEIVRMALFQDQNYTN-----PYEGDAVELSVESVLY 374
QY 186 AG-VEAKGLSRFKVLSNWCWATPSADPMYPLQWLKNGCPT--DET 242
DB 375 VGAILEQDTSRFLNLCVATPTEDKADLVKFIIRNSCSNQDSTI 434
QY 243 TFQFNAPRFQNIPLSKVWLHCETFCIDSEKLSCPVTCDK 282
DB 435 RFSVQMFAGHYDL--VFLHCEIHLCDLSNEQCPSCSR 472

RESULT 11
US-10-485-555-52
; Sequence 52, Application US/10485555
; Publication No. US20050120393A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Full Length Prostate Specific Polynucleotides and Polypeptides
; FILE REFERENCE: 9U 101.2 PCT
; CURRENT APPLICATION NUMBER: US/10/485,555
; CURRENT FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: US 60/309,470
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/330,747
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 52
; LENGTH: 534
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-485-555-52

Query Match
Best Local Similarity 12.0%; Score 210; DB 17; Length 534;
Matches 75; Conservative 43; Mismatches 104; Indels 58; Gaps 11;

QY 32 CYPKTIITIKIEPCYGVHQLGGLCYNGVHEGGYQFVLPDL-----76
DB 226 CGPREIKVKDKC-----LLGGL-----GLGEVIAYLRDPNCSILQTEERN 268
QY 77 -----SPKNKSYCGTQSEYKPIYHYFHYSHIVSNDTIVTKNQPNVNSFCTYHSTYLVN- 130
DB 269 WVSVTSPVQASACRNILE-RNQTHAIYKNTLSLVNDFIIRTNINQFQAYPLDMKVS 327
QY 131 QAADFQVRVATVHVKNVSGMTFESQLSL---NFYTNKFSIKKEAPFVLEASIG--SDLF 185
DB 328 QAALQPIVSSLSNVSDGNGEIVRMALFQDQNYTN-----PYEGDAVELSVESVLY 378
QY 186 AG-VEAKGLSRFKVLSNWCWATPSADPMYPLQWLKNGCPT--DET 242
DB 379 VGAILEQDTSRFLNLCVATPTEDKADLVKFIIRNSCSNQDSTI 438
QY 243 TFQFNAPRFQNIPLSKVWLHCETFCIDSEKLSCPVTCDK 282
DB 439 RFSVQMFAGHYDL--VFLHCEIHLCDLSNEQCPSCSR 476

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2005, 20:29:15 ; Search time 42 Seconds
(without alignments)
753.698 Million cell updates/sec

Title: US-10-687-268-35
Perfect score: 1756
Sequence: 1 MVTKAFVLLAIFAEASAKSC.....YSFSDVLHLLIMLIGICAVL 329

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1357.5	77.3	329	A57246	beta-tectorin prec
2	230	13.1	2155	T30197	alpha tectorin - m
3	228.5	13.0	644	A40212	uromodulin precurs
4	228.5	13.0	644	I84634	Tamm-Horsfall prot
5	222.5	12.7	642	S52111	uromodulin precurs
6	220	12.5	530	A38690	membrane glycoprot
7	219	12.5	640	A30452	uromodulin precurs
8	210	12.0	509	A37259	membrane glycoprot
9	210	12.0	530	G02091	pancreatic zymogen
10	204	11.6	2120	T30243	alpha tectorin - c
11	150.5	8.6	2083	T43721	CRP-ductin-alpha p
12	150.5	8.6	2403	A59386	sanko - human
13	148.5	8.5	540	S70403	zona pellucida gly
14	148.5	8.5	713	S70434	zona pellucida gly
15	147.5	8.4	1290	A57190	eberin precursor
16	147.5	8.4	1594	T30549	hensin - rabbit
17	144	8.2	713	A34782	sperm-binding gly
18	142	8.1	716	S70398	zona pellucida gly
19	134	7.6	745	A48833	sperm-binding gly
20	133	7.6	570	S70400	zona pellucida gly
21	130.5	7.4	715	S70397	zona pellucida gly
22	130	7.4	540	A39783	sperm-binding gly
23	120.5	6.9	841	I50428	transforming growt
24	117	6.7	536	S35712	sperm-binding gly
25	113.5	6.5	676	A45984	sperm-binding gly
26	111.5	6.3	918	TJ4361	scavenger receptor
27	99	5.6	928	VGBEMC	glycoprotein B pre
28	94	5.4	423	F86313	F2H15.8 protein -
29	94	5.4	1086	T43266	DNA-directed DNA p

30	94	5.4	1086	2	T40242	DNA polymerase del
31	92	5.2	853	1	A41220	transforming growt
32	91.5	5.2	887	2	A54832	villin homolog qua
33	90.5	5.2	1280	2	A39117	170K lectin precu
34	89.5	5.1	653	2	G96675	hypothetical prote
35	89.5	5.1	693	2	T19551	mucin-like protein
36	88.5	5.0	317	2	G59103	hypothetical prote
37	88.5	5.0	478	2	T02629	vacuolar processin
38	88.5	5.0	1084	1	S19661	DNA-directed DNA p
39	88.5	5.0	1292	2	T09229	galactose binding
40	88	5.0	698	2	H1535	hypothetical prote
41	87.5	5.0	399	2	D96717	hypothetical prote
42	87.5	5.0	400	2	T51953	H-protein promoter
43	87.5	5.0	837	2	A42112	mucin-like peptide
44	87	5.0	623	2	I46382	ZPI precursor - mo
45	86.5	4.9	440	2	T24323	hypothetical prote

ALIGNMENTS

RESULT 1

A57246
beta-tectorin precursor - chicken
C:Species: Gallus gallus (Chicken)
C>Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: A57246
R:Killick, R.; Legan, P.K.; Malenczak, C.; Richardson, G.P.
J. Cell Biol. 129, 535-547, 1995
A:Title: Molecular cloning of chick beta-tectorin, an extracellular matrix molecule of
A:Reference number: A57246; MUID:95238547; PMID:7721949
A:Accession: A57246
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-329 <KIL>
A:Cross-references: UNIPROT:P54097; GB:L38519; MID:g602439; PIDN:AAA32461.1; PID:g602440
A:Keywords: extracellular matrix; membrane protein; phosphatidylinositol linkage

Query Match 77.3%; Score 1357.5; DB 2; Length 329;
Best Local Similarity 73.6%; Pred. No. 4.9e-109;
Matches 242; Conservative 44; Mismatches 42; Indels 1; Gaps 1;

QY	1	MVTKAFVLLAIFAEASAKSCAPNKADVLVFCYPKTIIITKIPECPYGEVHQLALGLCY	60
DB	1	MVATVILMVTLAQAFGCTPNKADVLVFCYPTIITKIPECPYGEVHQLALGLCY	60
QY	61	NGVHEGGYQFVLPDLSPNKSYCGTQSEYKPIYHFYSHIVSNDTTVIVKNQPVNYSFS	120
DB	61	NGIHDSDGYQFTIPDLSPNKSYCGTQSEFKNPVHFYNSIVSNDSTVIVKSQPVNYSFT	120
QY	121	CTYHSTVLVNOAADFQDVATVHVKNSGMTFESQLSLNFYTNAPSIKKEAPFVLEASEI	180
DB	121	CTYNANYLVNOAADFQDVATVHVKNSGSGSPESQLSLNFYTNAPSIKKEAPFVLEASEI	180
QY	181	GSDLFAGVEAKGLSRFRKVLNSCWATPSADFMTPLOMLINKGCTPTDTVLVHENGDRH	240
DB	181	GSDIFAGVEAKGLSDRFKVLNWCWATPSSYFQVHWPLITKGCASDFSVIVHENGKTN	240
QY	241	RATQFNAPRFQNPILSKLKWHLHCETFTICDSEKLSQVTCDKRKLRLDQGGVLVAVLS	300
DB	241	RATQFNAPRFQNPILSKLKWHLHCETHVCDSEKFCPTVCDKRLQRM-EQTGGVLVAVIS	299
QY	301	LRSGFSLYSFSDVLHLLIMLIGICAVL	329
DB	300	VRNKGLSRFVMSDVIHLLFAIGFCAIL	328

RESULT 2

T30197
alpha tectorin - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30197

F:335-584/Domain: ZP domain homology <ZPH>

Query Match 12.7%; Score 222.5; DB 1; Length 642;
Best Local Similarity 26.3%; Pred. No. 4.3e-11;
Matches 55; Conservative 50; Mismatches 93; Indels 11; Gaps 7;

QY 76 LSPKNSYCGTQSEYKPIYHYSHIVSDTTVIYKQPNVYSFCTYHSTYLVN-QAAF 134
DB 382 VTPARNGPCGTVLK-RNETHATYNTLYLANAIIIRIIMNFECSYPLDMKVSLSKTS 440
QY 135 DQVATVHVKNKSGMTFESQSLNFTYNAKFIKEAPVLEASEGSLDFAGVAKGLS 194
DB 441 QPMVSAINSLGCTGKFTVRMAL--FQSPYTPQYQGPSVMLSTE--AFLYVGTMLDGGD 496

QY 195 I-RFKVVLNSCWATPSADPMYPLQWOLINKGCP--TDETVLVHNGRDHRAFFQFNAPRF 251
DB 497 LSRFVLLMNCVATPSNSTDPVKYFIQDSCPTREDTIQTEGESSQARFSVQMFAP 556
QY 252 QNPKLSKVLHCETFFICDSEKLSCPVTC 280
DB 557 AG--NYDLVYLHCEVYLCDSTSEQCKPTC 583

RESULT 6

A38690
membrane glycoprotein 2 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A38690; S11503; PNO573
R:Hoops, T.C.; Rindler, M.J.
J. Biol. Chem. 266, 4257-4263, 1991
A:Title: Isolation of the cDNA encoding glycoprotein-2 (GP-2), the major zymogen granule
A:Reference number: A38690; MUID:91154223; PMID:1999417
A:Accession: A38690
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-530 <HO>
R:Fukuoka, S.I.; Scheele, G.
Nucleic Acids Res. 18, 5900, 1990
A:Title: Nucleotide sequence encoding the major glycoprotein (GP2) of rat pancreatic sec
A:Reference number: S11503; MUID:91016950; PMID:2216794
A:Accession: S11503
A:Molecule type: DNA
A:Residues: 1-286, 'Q' 288-376, 'A' 378-530 <FUK>
A:Cross-References: EMBL:X53935; NID:957246; PIDN:CAA37882.1; PID:957247
R:Withiam-Leitch, M.; Aletha, J.M.; Koshlukova, S.E.; Rupp, G.; Beaudoin, A.R.; Rubin, R
Biochem. Biophys. Res. Commun. 194, 1167-1174, 1993
A:Title: Glycoprotein 2 of zymogen granule membranes shares immunological cross-reactivi
A:Reference number: PNO573; MUID:93356784; PMID:8352773
A:Accession: PNO573
A:Molecule type: protein
A:Residues: 327-376, 'A' 378-512 <WIT>
A:Experimental source: pancreas
C:Comment: This protein is a major protein of rat pancreatic zymogen granule membranes,
C:Superfamily: membrane glycoprotein 2; membrane glycoprotein 2 homology; ZP domain hom
C:Keywords: calcium binding; glycoprotein; lipoprotein; membrane protein; pancreas; phos
F:52-530/Domain: membrane glycoprotein 2 homology <MGH>
F:421-471/Domain: ZP domain homology <ZPH>
F:350,366,367,378/Binding site: phospholipid (Leu) #status predicted
F:375,377,397,420/Binding site: calcium (Val, Val, Thr, Asp) #status predicted

Query Match 12.5%; Score 220; DB 1; Length 530;
Best Local Similarity 27.1%; Pred. No. 5.5e-11;
Matches 78; Conservative 46; Mismatches 108; Indels 56; Gaps 12;

QY 23 NKADVI----LVFCYPTKIITIKIPCEPYGWEVHQLAGGLCYNVHGEGYQFVLPDLSP 78
DB 209 NVSDTSLQPLDCCGANEKVKLDKC-----LLGGL-----GFKEDIITYLND 251
QY 79 KNKSYC-GTQSEYKPIYHYSHIVSD-----TTVIVKQPNV 116
DB 252 RN--CRGTMKDEPNWVSTTSPVANDCGNLENGTHAIYRNTLSLATDPIIRDVLN 308

QY 117 YSFCTYHSTYLVN-QAAFQDQVATVHVKNKSGMTFESQSLNFTYNAKFIKEAPFVL 175
DB 309 VNFQCAYPDLMWVSLQTLQPIVSSLVNDVVGAGFTVTMAL--FQDQSYTHPYEGSKVL 366
QY 176 EASEIGSDFAGVEA-KGLSIRFKVVLNSCWATPSADPMYPLQWOLINKGCP--DETVL 232
DB 367 LPVE--NILYGVLLNRCGDSFRKLLTNCVATPSGDRNDIVKVFIRNRCPNQDSTIN 424
QY 233 VHEGRDHRATFQFNAPRFQNIPLKLSKVLHCETFFICDSEKLSCPVTC 280
DB 425 VEENGVSSESFSVQMFMPAG--NYDLVYLHCEVYLCDSTTEQCPSC 470

RESULT 7

A30452
uromodulin precursor - human
N:Alternate names: Tamm-Horsfall urinary glycoprotein; uromucoid
C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A30452; A30453
R:Pennica, D.; Kohr, W.J.; Kuang, W.J.; Glaister, D.; Aggarwal, B.B.; Chen, E.Y.; Goedde
Science 236, 83-88, 1987
A:Title: Identification of human uromodulin as the Tamm-Horsfall urinary glycoprotein.
A:Reference number: A30452; MUID:87177970; PMID:3453112
A:Accession: A30452
A:Molecule type: DNA; mRNA
A:Residues: 1-640 <PEN>
A:Cross-References: UNIPROT:P07911; GB:M15881; NID:9340163; PIDN:AAA36798.1; PID:9340164
A:Note: parts of this sequence, including the amino end of the mature protein, were date
R:Heslion, C.; Decker, J.M.; Sherblom, A.P.; Kumar, S.; Yue, C.C.; Mattaliano, R.J.; Tiza
Science 237, 1479-1484, 1987
A:Title: Uromodulin (Tamm-Horsfall glycoprotein): A renal ligand for lymphokines.
A:Reference number: A30453; MUID:87319675; PMID:3498215
A:Accession: A30453
A:Molecule type: mRNA
A:Residues: 1-640 <HES>
A:Cross-References: GB:M17778
A:Note: the authors translated the codon GTG for residue 381 as Asp
A:Note: parts of this sequence, including the amino end of the mature protein, were dete
R:Rindler, M.J.; Naik, S.S.; Li, N.; Hoops, T.C.; Peraldi, M.N.
J. Biol. Chem. 265, 20784-20789, 1990
A:Title: Uromodulin (Tamm-Horsfall Glycoprotein/Uromucoid) is a phosphatidylinositol-lin
A:Reference number: A30454; MUID:91065873; PMID:2249987
A:Contents: annotation; GPI-anchor
C:Comment: The protein is anchored to the cell membrane by a phosphatidylinositol linkag
e membrane-bound form.
C:Genetics:
A:Gene: GDB:UMOD
A:Cross-References: GDB:133729; OMIM:191845
A:Map position: 16p13.11-16p12.3
A:Introns: 30/1; 289/1; 325/1; 394/3; 444/2; 526/2; 580/3; 608/1; 621/1
A:Note: the first intron occurs before the initiator codon
C:Superfamily: uromodulin; BGF homology; membrane glycoprotein 2 homology; ZP domain hom
C:Keywords: duplication; glycoprotein; lipoprotein; membrane protein; phosphatidylinosit
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-640/Product: uromodulin #status predicted <MAT>
F:69-106/Domain: BGF homology <BGF>
F:170-639/Domain: membrane glycoprotein 2 homology <MGH>
F:334-583/Domain: ZP domain homology <ZPH>
F:38,76,80,275,322,513/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:232,396/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 12.5%; Score 219; DB 1; Length 640;
Best Local Similarity 25.7%; Pred. No. 8.5e-11;
Matches 66; Conservative 49; Mismatches 102; Indels 40; Gaps 9;

QY 76 LSPKNSYCGTQSEYKPIYHYSHIVSDTTVIYKQPNVYSFCTYHSTYLVN-QAAF 134
DB 381 VTPARDGCGVLT-RNETHATYNTLYLADEIIRDLNLIKINFACSYPLDMKVSLSKAL 439
QY 135 DQVATVHVKNKSGMTFESQSL-----NFTYNAKFIKEAPFVLEASIGSDFL 187

Db 440 QPMVSALNIRVGGTGMFTVRMALFQTPSYTQPYQSGSVTLSTEAFLYVGTMLDGGDL--- 496

Qy 188 VEAKGLSIRFKVKNLNSCWATPSADPMYPLQWLKNGCP--TDTVLVHNGRDHRAFPQ 245

Db 497 -----SRFALLMTCVATSSNATDPLKYFIIDRCPCPHTRDSTIQVNGESSQGRFS 549

Qy 246 FNAERFQNI PKL SKVWLHCETFCIDSEKLSCPVTCDEKRLLRDQTGGVL-----VVEL-- 299

Db 550 VQMFRRFAG--NYDLVYLHCETFCIDSEKLSCPVTCDEKRLLRDQTGGVL-----VVEL-- 299

Qy 300 -----SLRSGRFS 309

Db 604 ITRKGVQATVSRATSS 620

RESULT 8

membrane glycoprotein 2 precursor - dog

N:Alternate names: zymogen granule membrane associated protein ZAP75

C:Species: Canis lupus familiaris (dog)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: A37259; PC2195

R:Fukuoka, S.I.; Freedman, S.D.; Scheele, G.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 2898-2902, 1991

A:Title: A single gene encodes membrane-bound and free forms of GP-2, the major glycoprotein of zymogen granules in the dog

A:Reference number: A37259; MUID:91187898; PMID:2011597

A:Accession: A37259

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-509 <FUK>

A:Cross-references: UNIPROT:P25291; GB:M64083; GB:M37032; NID:g164102; PIDN:AAA30904.1; R:Fukuoka, S.

Biosci. Biotechnol. Biochem. 58, 1282-1285, 1994

A:Title: Analysis of ZAPs, zymogen granule membrane associated proteins, in the regulate

A:Reference number: PC2195; MUID:94362286; PMID:7765250

A:Accession: PC2195

A:Molecule type: protein

A:Residues: 22-32 <FUK>

A:Experimental source: pancreas

A:Note: This sequence is identical to that of glycoprophosphatidylinositol (GPI)-anchored p

C:Comment: This protein localizes in the zymogen granule membrane and has important func

C:Superfamily: membrane glycoprotein 2; membrane glycoprotein 2 homology; ZP domain hom

C:Keywords: calcium binding; glycoprotein; lipoprotein; membrane protein; pancreas; phos

F:34-509/Domain: membrane glycoprotein 2 homology <MGH>

F:200-450/Domain: ZP domain homology <ZPH>

Query Match 12.0%; Score 210; DB 1; Length 509;

Best Local Similarity 25.9%; Pred. No. 3.7e-10;

Matches 72; Conservative 43; Mismatches 115; Indels 48; Gaps 10;

Qy 32 CYPKTIITKIPCEPYGWEVHQLALGGLCYNGVHEGGYQFVIPDL----- 76

Db 201 CGAKEIQVSLDKC-----QLGGL-----GFGDEVIAIYLRDWNCSNMQRBERN 243

Qy 77 -----SPKNKSYCGTQSEYKPIYHFYSHIVNSDITVIVKNQPVNYSFCTVHSTYLVN- 130

Db 244 WISVTSPTQARACGNILE--RNQTHAIYKNTLSLVNDFIIRDITILNINFCAYPLDMKVS 302

Qy 131 QAAFDQVATVHVKNKSGMTFESQSLNFTYNKAFSIKKEAPFVLEASIGSDLPAG-VE 189

Db 303 QPALHPVSSLSNIVSDGEGFTVRMAL--FDQSYISPYEGAAVL---AVESMLVVGAIL 358

Qy 190 AKGLSIRFKVKNLNSCWATPSADPMYPLQWLKNGCP--TDTVLVHNGRDHRAFPQ 247

Db 359 EKGDTSRFLLRLNRYATPTKDTDPVKYFIIRNSCPNQYDSTIIVHENGVSSESFRFSVQ 418

Qy 248 AFRFQNI PKL SKVWLHCETFCIDSEKLSCPVTCDEKRLLRDQTGGVL-----VVEL-- 299

Db 419 MEFNAG--NYDLVYLHCETFCIDSEKLSCPVTCDEKRLLRDQTGGVL-----VVEL-- 299

RESULT 9

G02091

pancreatic zymogen granule membrane protein GP-2 - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 21-Jan-2000

C:Accession: G02091

R:Wong, S.M.E.; Lowe, A.W.

submitted to the EMBL Data Library, September 1995

A:Reference number: H00789

A:Accession: G02091

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-530 <WON>

A:Cross-references: EMBL:U36221; NID:g1244511; PIDN:AAB19240.1; PID:g1244512

C:Superfamily: membrane glycoprotein 2; membrane glycoprotein 2 homology; ZP domain hom

F:55-530/Domain: membrane glycoprotein 2 homology <MGH>

F:221-471/Domain: ZP domain homology <ZPH>

Query Match 12.0%; Score 210; DB 2; Length 530;

Best Local Similarity 26.8%; Pred. No. 4e-10;

Matches 75; Conservative 43; Mismatches 104; Indels 58; Gaps 11;

Qy 32 CYPKTIITKIPCEPYGWEVHQLALGGLCYNGVHEGGYQFVIPDL----- 76

Db 222 CGPREIKVYDKC-----LLGGL-----GLGEEVIAIYLRDNPNCSSILQTERN 264

Qy 77 -----SPKNKSYCGTQSEYKPIYHFYSHIVNSDITVIVKNQPVNYSFCTVHSTYLVN- 130

Db 265 WVSVTSPQASACRNILE--RNQTHAIYKNTLSLVNDFIIRDITILNINFCAYPLDMKVS 323

Qy 131 QAAFDQVATVHVKNKSGMTFESQSLNFTYNKAFSIKKEAPFVLEASIG--SDLP 185

Db 324 QAAQLQIVSSLSNIVSDGEGFTVRMALFDQNYTN-----PYEGDAVELSVESVLY 374

Qy 186 AG-VEAKGLSIRFKVKNLNSCWATPSADPMYPLQWLKNGCP--TDTVLVHNGRDHRA 242

Db 375 VGAILEQDTSRFLRLNRYATPTKDTDPVKYFIIRNSCPNQYDSTIIVHENGVSSES 434

Qy 243 TFQENAFRFQNI PKL SKVWLHCETFCIDSEKLSCPVTCDEKRLLRDQTGGVL-----VVEL-- 299

Db 435 RFSVQMFAGHYDL--VFLHCEIHLCDLSNEQCQPCSR 472

RESULT 10

alpha tectorin - chicken

C:Species: Gallus gallus (chicken)

C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004

C:Accession: T30243

R:Coutinho, P.; Goodyear, R.G.; Legan, P.K.; Richardson, G.P.

Hear. Res. 130, 62-74, 1999

A:Title: Chick alpha tectorin : molecular cloning and expression during embryogenesis.

A:Reference number: 220783; MUID:99251817; PMID:10320099

A:Accession: T30243

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2120 <COU>

A:Cross-references: UNIPROT:Q9VH85; EMBL:AJ012287; NID:e1361091; PID:e1361092; PIDN:CAAO

A:Note: non-collagenous protein only expressed in the inner ear

Query Match 11.6%; Score 204; DB 2; Length 2120;

Best Local Similarity 24.3%; Pred. No. 7.7e-09;

Matches 77; Conservative 48; Mismatches 128; Indels 64; Gaps 14;

Qy 30 VFCYKTIITKIPCEPYGWEVHQLALGGLCYNGVHEGGYQFVIPDLSPKNKSYCG- 85

Db 1771 VTCAAQMEVSIKCKLFGQFGREGVKNVDRHCPGIEGEDFISFQINN-----TKNGCN 1826

Qy 86 -----TQSEYKPIYHFYSHIVNSDITVIVKNQPVNYSFCTVHSTYLVNQAAPDQV- 138

Db 1827 LVQSNSTHIVYKNTV---WIESANNNGNIITRDRTINVEVFCAYE--LDIKISLDSVVR 1880

Qy 139 -----ATVHVKNKSGMTFESQSLNFTYNKAFSIK-KEAPFVLEASIGSDLPAGVEA 190

Db 1881 PMLSVINLTVPTQEGS---FTTKMAL--YKNSSYKHPYRQGEVLTTRDV---LVGVFV 1932

QY 191 KGL-SIRPKVILNSCWATPSADFMYPLOWOLINKGPT--DETVLVHNGRDRHATFQFN 247
Db 1933 VGADSNHLILMLANKYATPSRDSNDKLRYPFIIBGGQNLKNTIGIEENGVSUTCRFHYT 1992
QY 248 AFRFONIPKLSKVLMLHCETICDSEKLSCLSPVTCDKRKLRLRDQTGGVLVVELSLRSRGFS 307
Db 1993 VFKE--IGDYDEVHLHCAVSLCDSEKYSCKINCPQHR----- 2028
QY 308 SLYSFSDVLHHLIMLGG 324
Db 2029 SASAFAQEAHEQILSVG 2045
RESULT 11
T42721
CRP-ductin-alpha precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C/Accession: T42721
R/Cheng, H.; Bjerknes, M.; Chen, H.
Anat. Rec. 244, 327-343, 1996
A/Title: CRP-ductin: a gene expressed in intestinal crypts and in pancreatic and hepatic
A/Reference number: Z22241; MUID:96362470; PMID:8742698
A/Accession: T42721
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-2083 <CHE>
A/Cross-references: UNIPROT:Q60997; EMBL:U37438; NID:g1276646; PID:g1276647; PIDN:AAC525
A/Experimental source: strain BALB/c; jejunal epithelial cells
C/Keywords: transmembrane protein
F/1-28/Domain: signal sequence #status predicted <SIG>
F/29-2083/Product: CRP-ductin-alpha #status predicted <MAT>
Query Match 8.6%; Score 150.5; DB 2; Length 2083;
Best Local Similarity 24.0%; Pred. No. 0.0003;
Matches 47; Conservative 46; Mismatches 76; Indels 27; Gaps 10;
QY 84 CGT--QSEYKPIY-HFYSHIVSNDTIVVKGNQVNYFSC-----TYHST-YLVNQAAP 134
Db 1803 CGTIKQADNETINYSNPLRAWNSG--IIKRRKDLNIHVSCKMLQNTWNTYITNN--- 1857
QY 135 DORVATVHVKNGSMGTFESLSNIFYTNAKFSIK-KEAPFVLEASEIGSDLPAGVEAKGL 193
Db 1858 -----IVIEIQVQYGNFD--VNISFYTSSEFLFPVTSSPYV---DLQNLVLQAEILHS 1907
QY 194 SIRPKVILNSCWATPSADFMYPLOWOLINKGPTDETVLVHNGRDRHATFQFNAPRQFN 253
Db 1908 DASLALFVDTCVASPHNPFSSLTIDLIRSGCVRDDTYQSYSPSPRVRFRKFSFHFILN 1967
QY 254 IPKLSKVLMLHCETIFC 269
Db 1968 --RFPSPVILQCKLWVC 1981
RESULT 12
A59386
sanko - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-2001 #sequence_revision 31-Dec-2001 #text_change 09-Jul-2004
C/Accession: A59386
R/sanko, S.
submitted to the Protein Sequence Database, March 2001
A/Reference number: A59386
A/Accession: A59386
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-2403 <SAN>
A/Cross-references: UNIPROT:Q9UGM2
Query Match 8.6%; Score 150.5; DB 2; Length 2403;
Best Local Similarity 23.2%; Pred. No. 0.00036;
Matches 64; Conservative 47; Mismatches 102; Indels 63; Gaps 15;

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2005, 20:22:12 ; Search time 169 Seconds
(without alignments)
996.888 Million cell updates/sec

Title: US-10-687-268-35

Perfect score: 1756

Sequence: 1 MVTKAFVLLAFAEASAKS.....YSFSDVLLHLMGLGICAVL 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_treml.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1756	100.0	329	1	TECB HUMAN
2	1665	94.8	329	1	TECB MOUSE
3	1357.5	77.3	329	1	TECB_CHICK
4	251	14.3	415	2	Q95J36
5	250	14.2	431	2	Q8TCW7
6	247	14.1	415	2	Q8BGZ8
7	247	14.1	415	2	Q661R0
8	237	13.5	534	2	Q9D733
9	235.5	13.4	642	2	Q862Z3
10	230	13.1	2155	1	TECA MOUSE
11	229.5	13.1	573	2	Q9D6X7
12	228.5	13.0	644	1	UROM RAT
13	228.5	13.0	644	2	Q642D6
14	224	12.8	642	2	Q62285
15	223.5	12.7	642	2	Q91X17
16	221	12.6	2155	1	TECA HUMAN
17	220	12.5	643	1	UROM BOVIN
18	219	12.5	507	2	Q6ZS84
19	219	12.5	611	2	Q81YGO
20	219	12.5	640	1	UROM HUMAN
21	216	12.3	530	1	GP2 RAT
22	213	12.1	380	2	Q8NIA3
23	210	12.0	509	1	GP2 CANFA
24	210	12.0	527	1	GP2 HUMAN
25	210	12.0	530	2	Q6BD34
26	204	11.6	2120	1	TECA CHICK
27	202	11.5	797	2	Q8UW62
28	184.5	10.5	606	2	P70412
29	183	10.4	211	2	Q9CV56
30	182	10.4	413	2	Q8BYD5
31	181.5	10.3	607	2	O35360

32	181.5	10.3	607	2	Q9QZT0
33	178.5	10.2	1246	2	O6LA40
34	178.5	10.2	1374	2	O6LA41
35	178	10.1	429	2	Q91650
36	175.5	10.0	418	2	Q766V2
37	174	9.9	457	2	Q72660
38	172	9.8	375	2	Q8AWX2
39	172	9.8	457	2	Q72661
40	172	9.8	607	2	Q86UP6
41	171	9.7	272	2	Q8N216
42	171	9.7	326	2	Q86SG1
43	167	9.5	905	2	Q919R5
44	164	9.3	149	2	Q9CV52
45	164	9.3	357	2	Q9HAR7

Q9qzt0 rattus norv
Q6la40 homo sapien
Q6la41 homo sapien
Q91650 xenopus lae
Q766v2 gallus gall
Q72660 homo sapien
Q8awx2 xenopus tro
Q72661 homo sapien
Q86up6 homo sapien
Q8n216 homo sapien
Q86sg1 homo sapien
Q919r5 xenopus lae
Q9cv52 mus musculu
Q9har7 homo sapien

ALIGNMENTS

RESULT 1

TECB HUMAN
ID: Q96PL2; STANDARD; PRT; 329 AA.
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Beta-tectorin precursor.
GN Name=TECB;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Xia J.H., Zhang H.L., Deng H., Lu C.Y., Pang Q.;
RT "Cloning and identification of human TECB.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: One of the major non-collagenous components of the tectorial membrane (By similarity). The tectorial membrane is an extracellular matrix of the inner ear that covers the neuroepithelium of the cochlea and contacts the stereocilia bundles of specialized sensory hair cells. Sound induces movement of these hair cells relative to the tectorial membrane, deflects the stereocilia and leads to fluctuations in hair-cell membrane potential, transducing sound into electrical signals.
CC -!- SUBUNIT: May form homomeric filament after self-association or heteromeric filament after association with alpha-tectorin.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Probable). Found in the non-collagenous matrix of the tectorial membrane (By similarity).
CC -!- DOMAIN: Zona pellucida domain may enable to form filaments.
CC -!- PTM: The presence of a hydrophobic C-terminus preceded by a potential cleavage site strongly suggests that tectorins are synthesized as glycosylphosphatidylinositol-linked, membrane-bound precursors. Tectorins are targeted to the apical surface of the inner ear epithelia by the lipid and proteolytically released into the extracellular compartment.
CC -!- SIMILARITY: Contains 1 ZP domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

EMBL; AF312827; AAL12829.1; -;
Genew; HGNC:11721; TECB.
MIM; 602653; -;
InterPro; IPR001507; Endoglin/CD105.
Pfam; PF00100; Zona_pellucida; 1.

```

DR PRINTS: PRO0023; ZPELUCIDA.
DR SMART: SM00241; ZP: 1.
DR PROSITE: PS00682; ZP_DOMAIN; FALSE_NEG.
KW Extracellular matrix; Glycoprotein; GPI-anchor; Lipoprotein; Membrane;
FT SIGNAL.
FT CHAIN 1 17 Potential.
FT PROPEP 18 305 Beta-tectorin.
FT DOMAIN 306 329 Removed in mature form (Potential).
FT LIPID 19 283 ZP.
FT CARBOHYD 305 305 GPI-anchor amidated glycine (Potential).
FT CARBOHYD 80 80 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 104 104 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 116 116 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 145 145 N-linked (GlcNAc. .) (Potential).
FT SEQUENCE 329 AA; 36955 MW; 0B4D379624F48D0 CRC64;

Query Match 100.0%; Score 1756; DB 1; Length 329;
Best Local Similarity 100.0%; Pred. No. 5.9e-146;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTKAFVLLAIFAEASAKSCAPNKADVLVFCYPTKTIITKIPECPYGEVHQLALGLCY 60
DB 1 MVTKAFVLLAIFAEASAKSCAPNKADVLVFCYPTKTIITKIPECPYGEVHQLALGLCY 60
QY 61 NGVHEGGYQFVLPDLSPKNKSYCGTQSEYKPPYHFYSHIVSNDTIVVKNQPNVYSFS 120
DB 61 NGVHEGGYQFVLPDLSPKNKSYCGTQSEYKPPYHFYSHIVSNDTIVVKNQPNVYSFS 120
QY 121 CTHYSTYLVNQAAFDQVATVHVKNKSGMGTFFESQLSNFYTNNAKFSIKKEAPFVLESEI 180
DB 121 CTHYSTYLVNQAAFDQVATVHVKNKSGMGTFFESQLSNFYTNNAKFSIKKEAPFVLESEI 180
QY 181 GSDLPAGVEAKGLSTRFKVNLNSCWATPSADPMYPLQWLKNGCPTDVTVLVHNGRDH 240
DB 181 GSDLPAGVEAKGLSTRFKVNLNSCWATPSADPMYPLQWLKNGCPTDVTVLVHNGRDH 240
QY 241 RATQFNQAFRQNIPLKSKVWLHCETFFICDSEKLSKCPVTCDDRKLRLRDQTGGVLVVELS 300
DB 241 RATQFNQAFRQNIPLKSKVWLHCETFFICDSEKLSKCPVTCDDRKLRLRDQTGGVLVVELS 300
QY 301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329
DB 301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329

```

```

RESULT 2
TECB MOUSE STANDARD; PRT; 329 AA.
AC O08524;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Beta-tectorin precursor.
GN Name=Tectb;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Cochlea;
RX MEDLINE=97236843; PubMed=9079715; DOI=10.1074/jbc.272.13.8791;
RA Legan P.K., Rau A., Keene J.N., Richardson G.P.;
RT "The mouse tectorins. Modular matrix proteins of the inner ear
RT homologous to components of the sperm-egg adhesion system.";
RL J. Biol. Chem. 272:8791-8801(1997).
CC -!- FUNCTION: One of the major non-collagenous components of the
CC tectorial membrane (By similarity). The tectorial membrane is an
CC extracellular matrix of the inner ear that covers the
CC neuroepithelium of the cochlea and contacts the stereocilia
CC bundles of specialized sensory hair cells. Sound induces movement
CC of these hair cells relative to the tectorial membrane, deflects
CC the stereocilia and leads to fluctuations in hair-cell membrane

```

```

CC potential, transducing sound into electrical signals.
CC -!- SUBUNIT: May form homomeric filament after self-association or
CC heteromeric filament after association with alpha-tectorin.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Probable). Found in the non-collagenous matrix of the tectorial
CC membrane (By similarity).
CC -!- DOMAIN: Zona pellucida domain may enable to form filaments.
CC -!- PTM: The presence of a hydrophobic C-terminus preceded by a
CC potential cleavage site strongly suggests that tectorins are
CC synthesized as glycosylphosphatidylinositol-linked, membrane-bound
CC precursors. Tectorins are targeted to the apical surface of the
CC inner ear epithelia by the lipid and proteolytically released into
CC the extracellular compartment.
CC -!- SIMILARITY: Contains 1 ZP domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X99806; CAA68139.2; -.
CC MGD; MGI:109574; Tectb.
CC GO; GO:0005578; C:extracellular matrix; IDA.
CC GO; GO:0005201; F:extracellular matrix structural constituent; IDA.
CC InterPro; IPR001507; Endoglin/CD105.
CC Pfam; PF00100; Zona_pellucida; 1.
CC PRINTS; PR00023; ZPELUCIDA.
CC SMART; SM00241; ZP; 1.
CC PROSITE; PS00682; ZP_DOMAIN; FALSE_NEG.
KW Extracellular matrix; Glycoprotein; GPI-anchor; Lipoprotein; Membrane;
FT SIGNAL.
FT CHAIN 1 17 Potential.
FT PROPEP 18 305 Beta-tectorin.
FT DOMAIN 306 329 Removed in mature form (Potential).
FT LIPID 19 283 ZP.
FT CARBOHYD 305 305 GPI-anchor amidated alanine (Potential).
FT CARBOHYD 80 80 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 104 104 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 116 116 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 145 145 N-linked (GlcNAc. .) (Potential).
FT SEQUENCE 329 AA; 36985 MW; CC69A33BEC270AB3 CRC64;

Query Match 94.8%; Score 1665; DB 1; Length 329;
Best Local Similarity 94.2%; Pred. No. 5.9e-138;
Matches 310; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 MVTKAFVLLAIFAEASAKSCAPNKADVLVFCYPTKTIITKIPECPYGEVHQLALGLCY 60
DB 1 MVTKAFVLLAIFAEASAKSCAPNKADVLVFCYPTKTIITKIPECPYGEVHQLALGLCY 60
QY 61 NGVHEGGYQFVLPDLSPKNKSYCGTQSEYKPPYHFYSHIVSNDTIVVKNQPNVYSFS 120
DB 61 NGVHEGGYQFVLPDLSPKNKSYCGTQSEYKPPYHFYSHIVSNDTIVVKNQPNVYSFS 120
QY 121 CTHYSTYLVNQAAFDQVATVHVKNKSGMGTFFESQLSNFYTNNAKFSIKKEAPFVLESEI 180
DB 121 CTHYSTYLVNQAAFDQVATVHVKNKSGMGTFFESQLSNFYTNNAKFSIKKEAPFVLESEI 180
QY 181 GSDLPAGVEAKGLSTRFKVNLNSCWATPSADPMYPLQWLKNGCPTDVTVLVHNGRDH 240
DB 181 GSDLPAGVEAKGLSTRFKVNLNSCWATPSADPMYPLQWLKNGCPTDVTVLVHNGRDH 240
QY 241 RATQFNQAFRQNIPLKSKVWLHCETFFICDSEKLSKCPVTCDDRKLRLRDQTGGVLVVELS 300
DB 241 RATQFNQAFRQNIPLKSKVWLHCETFFICDSEKLSKCPVTCDDRKLRLRDQTGGVLVVELS 300
QY 301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329
DB 301 LRSRGFSSLSYFSDVLLHLLIMLIGICAVL 329

```

[illegible]

```
DR PRINTS: PR00023; ZPELUCIDA.
DR SMART; SM00241; ZP; 1.
KW Hypothetical protein.
SQ SEQUENCE 415 AA; 45487 MW; 9281DA2B3DBD8ED8ED CRC64;

Query Match
  14.3%; Score 251; DB 2; Length 415;
Best Local Similarity
  28.4%; Pred. No. 1.6e-13;
Matches 85; Conservative 46; Mismatches 134; Indels 34; Gaps 12;

QY 14 EASAKSCAPNKADVILVFCYPTKIITKIPECPY---GWEVHQLALGGL-----CYNGVHE 65
D 14 EASAKSCAPNKADVILVFCYPTKIITKIPECPY---GWEVHQLALGGL-----CYNGVHE 65
D 27 DANLHSRFPARD-IVSYGVQVAITWKINFCVTLVSGYSETDLALNGRHDGSHCRGFNN 85
QY 66 GGYQFVPIPLSPKNKSYCGTQ-----SEY-----KPIYHFYSHIVSNDTTIVKN 112
D 66 GGYQFVPIPLSPKNKSYCGTQ-----SEY-----KPIYHFYSHIVSNDTTIVKN 112
D 86 NTFPAVVIPIINLSTLEGCGNNLVSTIPGVSAYNATSVQGNISGYIDTDPPTIISY 145
QY 113 QP-VNYSFSCYTHSTVLVNOAFDQVATVHVKNKSGMTFFESQLSINFVYNAKFSIKKEA 171
D 113 QP-VNYSFSCYTHSTVLVNOAFDQVATVHVKNKSGMTFFESQLSINFVYNAKFSIKKEA 171
D 146 LPGLLYKFCSPLEYLEVNNVTQLASSAAISVRENN-GTFVSTLLNLLYNDSTYNOQLII 204
QY 172 PFVLEAETGSDLFAGVEAKGLSIRFKVVLNSCWATPSADFMYPQLQWLINKGCPDTEV 231
D 172 PFVLEAETGSDLFAGVEAKGLSIRFKVVLNSCWATPSADFMYPQLQWLINKGCPDTEV 231
D 205 PSI--GLPLTKVFAAQAATNLDGRWNVLMDCYTTTSGNPNDDIRYDLF--LSCDKDPQT 261
QY 232 LVHENGDRHRATQFNARFP--QNIPKLSKVLWLCETFCIDSEKLSCPV---TCDKRKR 285
D 232 LVHENGDRHRATQFNARFP--QNIPKLSKVLWLCETFCIDSEKLSCPV---TCDKRKR 285
D 262 TVIENGSRQGRFSFEVFRVFKHKNQKMTVFLHCVTKLCRAD--DCPFLMPCSHRER 318

RESULT 5
Q8TCW7 PRELIMINARY; PRT; 431 AA.
AC Q8TCW7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Zhu P.Y., Huang Y.F.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY090780; AAM09816.1;
DR InterPro; IPR001507; Endoglin/CD105.
DR Pfam; PF00100; Zona_pellucida; 1.
DR PRINTS; PR00023; ZPELUCIDA.
DR SMART; SM00241; ZP; 1.
KW Hypothetical protein.
SQ SEQUENCE 431 AA; 47377 MW; D7B0649DC9BAC073 CRC64;

Query Match
  14.2%; Score 250; DB 2; Length 431;
Best Local Similarity
  28.1%; Pred. No. 2.1e-13;
Matches 84; Conservative 47; Mismatches 134; Indels 34; Gaps 12;

QY 14 EASAKSCAPNKADVILVFCYPTKIITKIPECPY---GWEVHQLALGGL-----CYNGVHE 65
D 14 EASAKSCAPNKADVILVFCYPTKIITKIPECPY---GWEVHQLALGGL-----CYNGVHE 65
D 43 DANLHSRFPARD-IVSYGVQVAITWKINFCVTLVSGYSETDLALNGRHDGSHCRGFNN 101
QY 66 GGYQFVPIPLSPKNKSYCGTQ-----SEY-----KPIYHFYSHIVSNDTTIVKN 112
D 66 GGYQFVPIPLSPKNKSYCGTQ-----SEY-----KPIYHFYSHIVSNDTTIVKN 112
D 102 NTFPAVVIPIINLSTLEGCGNNLVSTIPGVSAYNATSVQGNISGYIDTDPPTIISY 161
QY 113 QP-VNYSFSCYTHSTVLVNOAFDQVATVHVKNKSGMTFFESQLSINFVYNAKFSIKKEA 171
D 113 QP-VNYSFSCYTHSTVLVNOAFDQVATVHVKNKSGMTFFESQLSINFVYNAKFSIKKEA 171
D 162 LPGLLYKFCSPLEYLEVNNVTQLASSAAISVRENN-GTFVSTLLNLLYNDSTYNOQLII 220
QY 172 PFVLEAETGSDLFAGVEAKGLSIRFKVVLNSCWATPSADFMYPQLQWLINKGCPDTEV 231
D 172 PFVLEAETGSDLFAGVEAKGLSIRFKVVLNSCWATPSADFMYPQLQWLINKGCPDTEV 231
```

```
Db 221 PSI--GLPLTKVFAAQAATNLDGRWNVLMDCYTTTSGNPNDDIRYDLF--LSCDKDPQT 277
Qy 232 LVHENGDRHRATQFNARFP--QNIPKLSKVLWLCETFCIDSEKLSCPV---TCDKRKR 285
Db 278 TVIENGSRQGRFSFEVFRVFKHKNQKMTVFLHCVTKLCRAD--DCPFLMPCSHRER 334

RESULT 6
Q8BGZ8 PRELIMINARY; PRT; 415 AA.
AC Q8BGZ8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Mus musculus 12 days embryo embryonic body between diaphragm region
DE and neck cDNA, RIKEN full-length enriched library, clone:9430016A21
DE product:hypothetical Endoglin / CD105 antigen containing protein, full
DE insert sequence (Mus musculus 16 days embryo head cDNA, RIKEN full-
DE length enriched library, clone:C130026H03 product:hypothetical
DE Endoglin / CD105 antigen containing protein, full insert sequence)
DE (Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched
DE library, clone:D630002M15 product:hypothetical Endoglin / CD105
DE antigen containing protein, full insert sequence).
GN Name=9430016A21Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Embryonic body between diaphragm region and neck, Head, and
RC Kidney;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Embryonic body between diaphragm region and neck, Head, and
RC Kidney;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Embryonic body between diaphragm region and neck, Head, and
RC Kidney;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Embryonic body between diaphragm region and neck, Head, and
RC Kidney;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
```


RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakauchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Ogawa Y., Izawa M., Ohara E., Watanabe J.,
RA Yoneda Y., Ishikawa T., Tanaka T., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Embryonic body between diaphragm region and neck, Head, and
RC Kidney;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saiton H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK034628; BAC28777.1; -;
DR EMBL; AK047981; BAC3205.1; -;
DR EMBL; AK052599; BAC35056.1; -;
DR MGD; MGI:2443415; 9430016A21Rik.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR001507; Endoglin/CD105.
DR Pfam; PF00100; Zona_pellucida; 1.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00241; ZP; 1.
KW Hypothetical protein.
SQ SEQUENCE 415 AA; 45447 MW; 5EDI871A4E257CA1 CRC64;

Query Match 14.1%; Score 247; DB 2; Length 415;
Best Local Similarity 27.8%; Pred. No. 3.7e-13;
Matches 83; Conservative 46; Mismatches 136; Indels 34; Gaps 11;

QY 14 ESASKCAPNKADVILVFCYKPTIITKIPECY---GWEVHOLALGGL-----CYNGVHE 65
Db 27 DANLHRSRFAERD-LSVCGVQAITWKINFCITVLSGYSETDLALNGRHDGSHCRGFNN 85
QY 66 GGYQYFVLPDLSPPKNSYCGTQ-----SEYKPIYHFYSHIVSNDTTIVKVN 112
Db 86 NTFFPAVVIIFINLSTLEGCGNNLVSTIPGVGASGNATTVOIGNISGYIDTDPDPAVISY 145
QY 113 QP-VNYSFSCCTHSTVYLNQAPDQRVATVHVKNVSGMGTFESQLSLNFYTNKAFSIKKEA 171
Db 146 LPGLLYKFCSCYPLEYLVNNTQLASSAAISVRENN-GTFVSTLNLLYNDSTYREQLII 204
QY 172 PFVLEASIGDLFAGVEAKGSIKPKVNLNSCWATPSADFMYQLQWLINKGCPDTEV 231
Db 205 PSI--GLPLKTKVFAVQATNLDGRWNVLMDYCYITPSPGNPNDDTRIDYLF-LSCDKDPOT 261
QY 232 LVHENGDRHRAVQFNAFRF--QNTIPKLSKWLHCETFCDSKLSKCPV---TCDKRRK 285
Db 262 TWIENGSRGRGFSFEVFFVFKHKNQKMTVFLHCLTKLCRAD--DCPLLMPICGNRRK 318

RESULT 7
Q66IR0 PRELIMINARY; PRT; 415 AA.
AC Q66IR0;
DT 25-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DE MGC85537 protein.
GN Name=MGC85537;

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8935;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX PubMed=12477992; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Atlachul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fallaj J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RN SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Klein S., Gerhardt D.S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC081237; AAH81237.1; -;
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR006162; Ppantne_S.
DR Pfam; PF00100; TONB_Box_N.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; UNKNOWN 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN 1.
SQ SEQUENCE 415 AA; 45730 MW; 3787133B98507169 CRC64;

Query Match 14.1%; Score 247; DB 2; Length 415;
Best Local Similarity 27.5%; Pred. No. 3.7e-13;
Matches 88; Conservative 47; Mismatches 135; Indels 50; Gaps 13;

QY 7 VLLAIF-----AESAKSCAPNKADVILVFCYKPTIITKIPECY---GWRVH 51
Db 8 LLLAIFTVSVAQFNQNCANDANQHSRFAERD-ITVYGVQITWKINFCITVLSGYSES 66
QY 52 QIALGGL-----CYNGVHEGYQYFVLPDLSPPKNSYC-----GTQSYK 91
Db 67 DLSLNGKHGDAHCRGFNNNTFPVITFNTLSTLESCNSLVSVTVPGVNAYGATSMVQ 126
QY 92 PPIVHFYSHIVSNDTTIVKQNQ--VNYSFSCCTHSTVYLNQAPDQRVATVHVKNVSGMT 150
Db 127 --IGNISGYIDTDPPTIISYLPGLLYKFCSCYPLEYLVNNTQLASSAAISVREGN-GT 183
QY 151 FESQLSLNFYTNKAFSIKKEAPFVLEASIGDLFAGVEAKGSLIRFKVNLNSCWATPSA 210
Db 184 FISTLNLLYNDSTYSQQLLIPSA--GLPLTKIYAAVRATNLDGRWNVLMDYCYITPFG 241


```

DR InterPro: IPR001507; Endoglin/CD105.
DR Pfam: PF07645; EGF CA; 2.
DR Pfam: PF00100; Zona_pellucida; 1.
DR PRINTS: PR00023; ZPELLUCIDA.
DR SMART: SM00179; EGF_CA; 2.
DR SMART: SM00241; ZP; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS00026; EGF_3; 3.
DR PROSITE: PS01187; EGF_CA; 2.
DR PROSITE: PS00682; ZP_DOMAIN; 1.
DR EGF-like domain.
SQ SEQUENCE 642 AA; 70177 MW; 608ACCF13A667E64 CRC64;

Query Match
Best Local Similarity 13.4%; Score 235.5; DB 2; Length 642;
Matches 84; Conservative 62; Mismatches 140; Indels 57; Gaps 12;

QY 2 VTKAFVLLAFAEASAKSCAPNKADVILVFCYPKTIITKIPECPYQWGVHOLALGGLCVN 61
DB 325 VTDLFLL-----DRLECRPNIDKVLKSKQLKSL-----GPEKVFMYLRDSQCS 368

QY 62 GVHEGGYQFVLPIDLSPRKNSVCGTQSEYKPIYHYFVSHVNSDITVIVKNQPVNYSFSC 121
DB 369 GFNERGDRDW-SVTPARDGPGCTVM-VRNETHATYSNTLYLADBIIVDRNINKEFEC 426

QY 122 TYHSTYLVN-QAFOFQVATVHVKNMGCTPESQLSL-----NFTYNAKESIKKEAPF 173
DB 427 SYPLDMKVSLETSLOPIVSSLSINVGVTGMFTVRMALFQTPDYTPQYSSVTLTLEAFL 486

QY 174 VLEASIEGDLFAGVEAKGLSIRFKVLNLSWCATPSADPMYFLOWOLINKGCP--TDFTV 231
DB 487 YGTWLDGGDL-----SRFALLMNCVATPSSNATDPLKFIIDRCPRITDSTI 536

QY 232 LVHENGDRHATFQFNAPFQNIPLSKVWLHCETFCIDSEKSLCPVTCREKRLLRDQT 291
DB 537 QVVENGESQGRFSVQMFREAG--NYDLVYLHCEVVLCDIINEKCKPTCSG---TRFRS 590

QY 292 GGVV-----VVEL-----SLRGRGSSLSYFSDVLHLIM 321
DB 591 GGIIDQSRVNLNIGPITRKNQVAVSRASSSIGFLKVLCLPLL 633

RESULT 10
TECA MOUSE
ID TECA MOUSE STANDARD; PRT; 2155 AA.
AC O08523;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Alpha-tectorin precursor.
GN Name=Tecta;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), SEQUENCE OF 25-34, SUBUNITS,
RP SUBCELLULAR LOCATION, POST-TRANSLATIONAL MODIFICATIONS, AND TISSUE
RP SPECIFICITY.
RC STRAIN=CD-1; TISSUE=Cochlea;
RX MEDLINE=97236843; PubMed=9079715; DOI=10.1074/jbc.272.13.8791;
RA Legan P.K., Rau A., Keene J.N., Richardson G.P.;
RT "The mouse tectorins. Modular matrix proteins of the inner ear
RT homologous to components of the sperm-egg adhesion system.";
RL J. Biol. Chem. 272:8791-8801(1997).
RN [2]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP MEDLINE=98250172; PubMed=9590290;
RX Verhoeven K., Van Laer L., Kirschhofer K., Legan P.K., Hughes D.C.,
RA Schattman I., Verstrecken M., Van Hauwe P., Coucke P., Chen A.,
RA Smith R.J.H., Somers T., Offeciers F.E., Van de Heyning P.,

```

```

RA Richardson G.P., Wachtler F., Kimberling W.J., Willems P.J.,
RA Govaerts P.J., Van Camp G.;
RT "Mutations in the human alpha-tectorin gene cause autosomal dominant
RT non-syndromic hearing impairment.";
RL Nat. Genet. 19:60-62(1998).
CC -!- FUNCTION: One of the major non-collagenous components of the
CC tectorial membrane (By similarity). The tectorial membrane is an
CC extracellular matrix of the inner ear that covers the
CC neuroepithelium of the cochlea and contacts the stereocilia
CC bundles of specialized sensory hair cells. Sound induces movement
CC of these hair cells relative to the tectorial membrane, deflects
CC the stereocilia and leads to fluctuations in hair-cell membrane
CC potential, transducing sound into electrical signals.
CC -!- SUBUNIT: May form homomeric filament after self-association or
CC heteromeric filament after association with beta-tectorin.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Probable). Found in the non-collagenous matrix of the tectorial
CC membrane.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O08523-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O08523-2; Sequence=VSP_010557;
CC -!- TISSUE SPECIFICITY: Cochlea-specific.
CC -!- DOMAIN: Zona pellucida domain may enable to form filaments.
CC -!- PTM: 3 products of tectorin seem to exist: HMM, MMM and LMM. They
CC may be generated by active processing or the result of proteolysis
CC occurring between intrachain disulfid bonds.
CC -!- PTM: The presence of a hydrophobic C-terminus preceded by a
CC potential cleavage site strongly suggests that tectorins are
CC synthesized as glycosylphosphatidylinositol-linked, membrane-bound
CC precursors. Tectorins are targeted to the apical surface of the
CC inner ear epithelia by the lipid and proteolytically released into
CC the extracellular compartment.
CC -!- SIMILARITY: Contains 1 NIDO (nidogen-like) domain.
CC -!- SIMILARITY: Contains 3 TIL (Trypsin inhibitory-like) domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 4 VWFD domains.
CC -!- SIMILARITY: Contains 1 ZP domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X99805; CAA68138.1; -.
CC PIR; T30197; T30197.
CC HSSP; P56682; ICCV.
CC MGD; MG1:109575; Tecta.
CC GO; GO:0005578; C:extracellular matrix; IDA.
CC GO; GO:0005201; F:extracellular matrix structural constituent; IDA.
CC InterPro; IPR002086; Aldehyde dehydr.
CC InterPro; IPR002919; Cysrich_TIL.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR001507; Endoglin/CD105.
CC InterPro; IPR000421; FA58_C.
CC InterPro; IPR006210; IEGF-.
CC InterPro; IPR010362; NIDO.
CC InterPro; IPR003886; Nidogen_ext.
CC InterPro; IPR001007; VWF_C.
CC InterPro; IPR001846; VWF_D.
CC Pfam; PF06119; NIDO; 1.
CC Pfam; PF01826; TIL; 3.
CC Pfam; PF00094; VWD; 4.
CC Pfam; PF00100; Zona_pellucida; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00539; NIDO; 1.
CC SMART; SM00214; VWC; 1.
CC SMART; SM00216; VWD; 4.

```

DR SMART: SM00241; 2P: 1.
 DR PROSITE, PS00184; VWFC 2; FALSE_NEG.
 DR PROSITE, PS00682; ZP_DOMAIN; 1.
 KW Alternative splicing; Direct protein sequencing; Extracellular matrix;
 KW Glycoprotein; GPI-anchor; Lipoprotein; Membrane; Repeat; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 2091
 FT PROPEP 2092 2155
 FT LIPID 2091 2091
 FT DOMAIN 98 254
 FT CARBOHYD 34 34
 FT CARBOHYD 187 215
 FT CARBOHYD 215 215
 FT CARBOHYD 278 278
 FT CARBOHYD 455 455
 FT CARBOHYD 506 506
 FT CARBOHYD 528 528
 FT CARBOHYD 560 560
 FT CARBOHYD 670 670
 FT CARBOHYD 687 687
 FT CARBOHYD 813 813
 FT CARBOHYD 843 843
 FT CARBOHYD 855 855
 FT CARBOHYD 898 898
 FT CARBOHYD 920 920
 FT CARBOHYD 931 931
 FT CARBOHYD 949 949
 FT CARBOHYD 1048 1048
 FT CARBOHYD 1064 1064
 FT CARBOHYD 1235 1235
 FT CARBOHYD 1364 1364
 FT CARBOHYD 1538 1538
 FT CARBOHYD 1565 1565
 FT CARBOHYD 1756 1756
 FT CARBOHYD 1772 1772
 FT CARBOHYD 1794 1794
 FT CARBOHYD 1851 1851
 FT CARBOHYD 1864 1864
 FT CARBOHYD 1880 1880
 FT CARBOHYD 1920 1920
 FT CARBOHYD 1939 1939
 FT VARSPLIC 1659 1663
 SQ SEQUENCE 2155 AA; 239535 MW; 86C5C09AA392B1C5 CRC64;
 Query Match 13.1%; Score 230; DB 1; Length 2155;
 Best Local Similarity 27.5%; Pred. No. 8.8e-11;
 Matches 80; Conservative 43; Mismatches 124; Indels 44; Gaps 14;
 QY 23 NKADVI--LVFCVPKTIITKIPEC---PYGWEVHOLALGGLCYNGHEGGYQFVIPDL 77
 1795 NSHDIIDAEVTKCAQMEVISKCKLFQGFREGVRINDRQCSGIEGEDFISFOINN-- 1852
 QY 78 PKNKSYC-----GTQSYKPPPIVHYFVSHIVNSDTTIVKNQPNVYGFSCYHTHYLVN 130
 1853 --TKGNCGNIVQSNTHMYKNTI---WIESANNTGNIITRDITINVEFSAYE---LDI 1904
 QY 131 QAPDORV-----ATVHKVNGSMGTPESQSLNFTYNKFSIK-KEAPTVLEASEIG 181
 1905 KISLDSVVKPMLSVINLTVPTQEGS---FTTKVAL--YKNASYKHPIYQGEVVLTRDV- 1958
 QY 182 SOLFAGVEAKGL-SIRFKVVLNSCWATPSADPFMYPLQWOLINKGCPT--DETVLVHNGR 238

Db 1959 --LYVGVVVGADSTHLLTLNKCYATPSSRDSNKLRYFIEGGCQCNKNTIGIEENG 2016
 QY 239 DHRATFOFNFRQNPISKLVMLHCETFTICDSEKLSCTVTCDKRKLLED 289
 Db 2017 SLTCRFHVTVFKF--IGDYDEVHLHCAVSLCDSEKYSKINCPCQNSRIATD 2065
 RESULT 11
 Q9D6X7
 ID Q9D6X7 PRELIMINARY; PRT; 573 AA.
 AC Q9D6X7
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus adult male tongue cDNA, RIKEN full-length enriched
 DE library, clone:2310046E09 product:membrane glycoprotein 2
 DE homolog.
 GN Name=Gp2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Tongue;
 MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Tongue;
 MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Tongue;
 MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA The FANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [4]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Tongue;
 MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 Kanno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Tongue;
 MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 Kanno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 Onoda Y., Ishikawa T., Orawa K., Tanaka T., Matsura S., Kawai J.,
 Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Tongue;
 ADACHI J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,


```
Db 384 VTPARDGCGTVLR--RNETHATYSNTLYLASEIIRDIRINIRINFECSYPLDMKVS�KTSL 442
Qy 135 DORVATVHVKNKSGMTFESQSLNFTYNAKFSIKKEAPFVLEASIGSDLPAGVEAKGLS 194
Db 443 QPMVSALNISLGGTGKFTVQMAL--FQNPPTYQPYQGPSVLMSTE--AFLYVGTMLDGGD 498
Qy 195 I-RFKVVLNSCWATPSADPMYPLQWLINKGCP--TDETVLVHNGRDRHRTAFQNAFRF 251
Db 499 LSRFVLLMTCYATPSSNSTDPVKYFIQDRCPHTEDTTIQVTENGESSQARFSIQMFRF 558

Qy 252 QNIPKLSKVLHCETFCIDSEKLSCPVTC 280
Db 559 AGNSDL--VYLHCEVYLCDTMSQCKPTC 585

RESULT 13
Q642D6 PRELIMINARY; PRT; 644 AA.
AC Q642D6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Uromodulin.
GN Name=Umod;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DE EMBL; BC081814; AAH81814.1; -.
SQ SEQUENCE 644 AA; 71016 MW; 2414261D3472ADEC CRC64;

Query Match 13.0%; Score 228.5; DB 2; Length 644;
Best Local Similarity 27.3%; Pred. No. 2.7e-11;
Matches 57; Conservative 50; Mismatches 91; Indels 11; Gaps 7;

Qy 76 LSPKNKSYCGTQSEYKPPHYHYSHIVSNDDTVIVKNQPNVYFSCTVHSTYLVN-QAAF 134
Db 384 VTPARDGCGTVLR--RNETHATYSNTLYLASEIIRDIRINIRINFECSYPLDMKVS�KTSL 442
Qy 135 DORVATVHVKNKSGMTFESQSLNFTYNAKFSIKKEAPFVLEASIGSDLPAGVEAKGLS 194
Db 443 QPMVSALNISLGGTGKFTVQMAL--FQNPPTYQPYQGPSVLMSTE--AFLYVGTMLDGGD 498
Qy 195 I-RFKVVLNSCWATPSADPMYPLQWLINKGCP--TDETVLVHNGRDRHRTAFQNAFRF 251
```

```
Db 499 LSRFVLLMTCYATPSSNSTDPVKYFIQDRCPHTEDTTIQVTENGESSQARFSIQMFRF 558
Qy 252 QNIPKLSKVLHCETFCIDSEKLSCPVTC 280
Db 559 AGNSDL--VYLHCEVYLCDTMSQCKPTC 585

RESULT 14
Q62285 PRELIMINARY; PRT; 642 AA.
AC Q62285;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Uromodulin.
GN Name=Umod; Synonym=Tamm-Horsfall protein;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95178555; PubMed=7873609; DOI=10.1016/0167-4781(94)00240-4;
RX Prasadani K., Bates J., Badgett A., Dell M., Sukhatme V., Yu H.,
RX Kumar S.;
RT "Nucleotide sequence and peptide motifs of mouse uromodulin (Tamm-
RT Horsfall protein)-the most abundant protein in mammalian urine."
RL Biochim. Biophys. Acta 1260:328-332 (1995).
DR EMBL; L33406; AAA73896.1; -.
DR PIR; S52111; S52111.
DR HSSP; P35555; 1EMN.
DR MGD; MGI:102674; Umod.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001507; Endoglin/CD105.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF07645; EGF_CA; 2.
DR Pfam; PF00100; Zona_pellucida; 1.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS00026; EGF_3; 3.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW EGF-like domain.
SQ SEQUENCE 642 AA; 70839 MW; 415A7E58D3D2B1A9 CRC64;

Query Match 12.8%; Score 224; DB 2; Length 642;
Best Local Similarity 25.5%; Pred. No. 6.6e-11;
Matches 65; Conservative 58; Mismatches 118; Indels 14; Gaps 8;

Qy 76 LSPKNKSYCGTQSEYKPPHYHYSHIVSNDDTVIVKNQPNVYFSCTVHSTYLVN-QAAF 134
Db 384 VTPARDGCGTVLR--RNETHATYSNTLYLASEIIRDIRINIRINFECSYPLDMKVS�KTSL 440
Qy 135 DORVATVHVKNKSGMTFESQSLNFTYNAKFSIKKEAPFVLEASIGSDLPAGVEAKGLS 194
Db 441 QPMVSALNISLGGTGKFTVQMAL--FQNPPTYQPYQGPSVLMSTE--AFLYVGTMLDGGD 496
Qy 195 I-RFKVVLNSCWATPSADPMYPLQWLINKGCP--TDETVLVHNGRDRHRTAFQNAFRF 251
Db 497 LSRFVLLMTCYATPSSNSTDPVKYFIQDRCPHTEDTTIQVTENGESSQARFSIQMFRF 556
Qy 252 QNIPKLSKVLHCETFCIDSEKLSCPVTC--RKRLLRDQTGGVLVVELSLSRGFS 308
```

```
Db 557 AG-NYDLVYLCEVYLCDSSTSEQCKPTCGTRFCGNFIDQTRVNLGPIRQRGVQA 614
Qy 309 LYSFSDVLHLLIMWL 323
Db 615 SVSASSNRLLSIWL 629

RESULT 15
Q91X17
ID Q91X17 PRELIMINARY; PRT; 642 AA.
AC Q91X17;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)
DE Uromodulin (Mus musculus 0 day neonate kidney cDNA, RIKEN full-length
DE enriched library, clone:D630023M23 product:uromodulin, full insert
DE sequence).
GN Name=Umod;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver; DOI=10.1073/pnas.242603899;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Ziegler B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
```

```
RL Nature 420:563-573(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirao T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takagawa A., Takada Y., Tanaka T.,
RA Tonaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC012973; AAI12973.1; -
DR EMBL; AK085460; BAC39452.1; -
DR PIR; S52111; S52111.
DR HSSP; P35555; 1EMN.
DR MGD; MGI:102674; Umod.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR Pfam; PF07645; EGF_1.
DR Pfam; PF00100; Zona pellucida; 1.
DR PRINTS; PR00023; ZPELUCLIDA.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00010; ASX HYDROXYL; 2.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS00026; EGF_3; 3.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
DR EGF-like domain.
SQ SEQUENCE 642 AA; 70844 MW; 31B1461B4DCAE927 CRC64;
```

```
Query Match 12.7%; Score 223.5; DB 2; Length 642;
Best Local Similarity 26.3%; Pred. No. 7.3e-11;
Matches 55; Conservative 50; Mismatches 93; Indels 11; Gaps 7;

Oy 76 LSPKKNKSYCGTQSEYKPPYHYSHIVNSDITVVKQPNVNSFCTVHSTYLVN-QAAF 134
Db 382 VTPARNGFCGTGLR-RNETHATYNTLYLANAIIRIIRIMNFCYPLDMKVLKLSL 440
Oy 135 DQRVATVHKNKSGMTFESQLSNFYTNKAPFKSEAPFVLEASEISGDLFAGVEAKGLS 194
Db 441 QPMYSALNISLGGTGKFTVRMAL--FQSTVTYQPHGQPSVMLSTE--AFLVGTMLDGGD 496
```

Qy	195	I-REKVLNSCWATPSADFMYPLOWOLINKGCP--TDETVLVHENGDRHRATQFNAPRF	251
Db	497	LSRFVLLMNCYATPSSNSTDPVKFILODSCPRTEDTIQVTENGESSQARFSVQMF	556
Qy	252	QNIPLKLSKVWLHCETFI CDSEKLSCPVTC	280
Db	557	AG--NYDLVYLHCEVYLC DSTSEQCKPTC	583

Search completed: July 20, 2005, 20:37:07
Job time : 172 secs